HOST DERIVED MARKERS OF LYME DISEASE AND OTHER SPIROCHAETAL INFECTIONS: THEIR DISCOVERY AND DIAGNOSTIC POTENTIAL

Greg Philip Joyner

Thesis submitted in accordance with the

requirements of the University of Liverpool

for the degree of Doctor of Philosophy (Ph.D.)

February 2020

ACKNOWLEDGEMENT

This research was funded by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Emerging and Zoonotic Infections at the University of Liverpool in partnership with Public Health England and Liverpool School of Tropical Medicine.

Many thanks to my supervisors; Dr Nick Beeching, Dr Amanda Semper, Prof Julian Hiscox and Dr Tim Brooks. Thanks to Dr Stuart Armstrong for all his help with my project. Thank you to all the staff at the RIPL who helped me conduct Lyme disease testing, and for providing samples for this study. Thanks to Dr Ales Chrdle for providing whole-blood samples from the Czech Republic. Thanks to Dr Nicholas Evans and Dr Roger Evans who conducted my viva.

Thank you to all the members of the Hiscox research group, It was a joy to be part of.

A final and sincere thank you to all the new friends I made in Liverpool, to my mum and dad, and to Mahatma Ghandi.

ABSTRACT

HOST DERIVED MARKERS OF LYME DISEASE AND OTHER SPIROCHAETAL INFECTIONS: THEIR DISCOVERY AND DIAGNOSTIC POTENTIAL

UK laboratory diagnosis of Lyme disease involves the Standard Two-Tier serological approach. The negative predictive value of the test has been challenged, particularly in the early stages of disease. In order to better understand the host response to Lyme disease and to identify any potential host biomarkers that may correlate with early infection, proteomic and transcriptomic analyses were undertaken on Lyme disease patient samples.

Label free quantitative spectrometry was used to measure and compare the serum proteome of seropositive patients against that of individuals that had tested seronegative for Lyme disease and a control group consisting of samples from normal healthy donors and patients with related infectious disease including leptospirosis and syphilis. Seropositive and seronegative individuals were found to be remarkably similar at the serum proteome level. Of the 12 proteins found to be at significantly different abundance between groups, the protein Lipocalin-2 was of particular interest due to role its role modulation of immune responses. Further analysis by ELISA in additional samples showed that Lipocalin-2 was significantly increased in Lyme disease positive patients when compared to normal healthy donors.

Whole blood samples from patients with Lyme disease were then RNA sequenced and differential gene analysis was performed using sequencing data for healthy controls. Several pathways associated with bacterial immune response in the host were

identified, together with actin rearrangement and oxidative stress pathways. eIF2 signalling was found to be significantly reduced in patients with early Lyme disease. The pathway has previously been identified as being down regulated during Lyme disease. In later samples from patients, following antibiotic therapy, eIF2 signalling levels were found to increase back towards those seen in control samples from normal/healthy individuals.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACA	Acrodermatitis chronicum atrophicans
AKI	Acute kidney injury
APS	Ammonium persulfate
B-ME	beta-mercaptoethanol
BCA	Bicinchoninic acid assay
CDC	Centers for Disease Control and Prevention
CLD	Chronic Lyme Disease
CNS	Central nervous system
DEG	Differentially expressed gene
DGE	Differential gene expression
EGFR	Epidermal growth factor receptor
EIA	Enzyme-linked assay
ELISA	Enzyme-linked immunosorbent assay
EBOV	Ebola virus
EM	Erythema migrans
FC	Fold change
HRP	Horseradish peroxidase
IFN	Interferon
Ig	Immunoglobulin
IL	Interleukin
LA	Lyme arthritis
LD	Lyme disease
LLD	Late Lyme disease
LTT	Lymphocyte transformation test
LNB	Lyme neuroborreliosis
MAT	Microscopic Agglutination Test
MS/MS	Tandem mass spectrometry
NICE	National Institute for Health and Care Excellence
NHS	National Health Service
NTT	Nontreponemal test

OSP	Outer surface protein
PBMC	Peripheral blood mononuclear cel
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PHE	Public Health England
PrEP	Pre-exposure prophylaxis medication
PTLDS	Post-treatment Lyme disease syndrome
RDT	Rapid diagnostic test
RIPL	Rare and Imported Pathogens Lab
RNA	Ribonucleic acid
RT-PCR	Real time polymerase chain reaction
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
STT	Standard two-tier (testing)
s.l.	Sensu lato
s.s.	Sensu stricto
TBS-T	Tris buffered saline - tween 20
THPA	*T. pallidum* hemagglutination assay
TMB	tetramethylbenzidine
TNF	Tumour necrosis factor

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

The main aim of this thesis was to use proteomic and transcriptomic approaches to investigate the blood of Lyme disease patients and to identify host protein and/or gene biomarkers that may have future diagnostic or prognostic potential. The majority of Lyme disease diagnosis currently relies on the detection of host antibodies to *Borrelia burgdorferi* sensu lato (s.l.) antigens in patient serum. Particularly in early disease, when the patient is mounting an immune response to infection, the negative predicative value of these tests has been questioned and there is therefore great interest in the development of improved diagnostic methods. Here, a background on Lyme disease including an in-depth discussion of the current diagnostic methods is given together with a review of the use of 'omics-based technologies in biomarker discovery studies.

During mass spectrometry analyses in chapter 3, serum samples from leptospirosis and syphilis positive patients were included as a related-disease control group. A range of differentially abundant proteins were observed in these groups and some protein markers were taken forward for further analyses in chapter 4, both as a proof-of-concept of the proteomic approach and because they represented interesting and novel findings; Therefore, background information regarding these spirochaetal infectious diseases is also included in this chapter.

1.1. CLASSIFICATION, DISCOVERY AND EPIDEMIOLOGY

1.1.1. The Spirochaete phylum

A spirochaete is a member of the phylum Spirochaetes. Spirochaetes are long, helically coiled, highly motile bacteria and can be described as gram-negative as they possess inner and outer membranes separated by a peptidoglycan layer within a periplasmic space. The bacteria are generally from 3 to 250 μm in length and 0.1-75 μm in diameter (1). The spirochaetes have a unique cellular structure, with flagellae located internally within the periplasmic space. The specialised flagella, termed axial filaments, rotate within the periplasm causing propulsion of the bacterial cell. The phenotypic characteristics shared between species of the phyla reflect the phylogenetic relationships between genera (2). 16S rRNA gene sequence comparisons show a distinct evolutionary progression within the phyla. **Figure 1.1** shows the phylogenetic tree of the spirochaetes. Species within several genera are free-living in the environment and anaerobic, but the group is diverse and there are several exceptions. Beyond the presence of the distinctive endoflagella, genera within the species share few other characteristics and are widespread in the environment. Despite being a relatively small phylum of bacteria, several species cause prevalent disease in humans. Clinically important genera include the *Borrelia, Treponema, Leptospira* and *Brachyspira*. **Figure 1.2** shows a phylogenetic tree based on protein homology between genospecies of the *Borrelia* genus. The tree includes data from 30 *Borrelia* strains including 17 associated with Lyme disease/borreliosis and 11 associated with relapsing fevers (3).

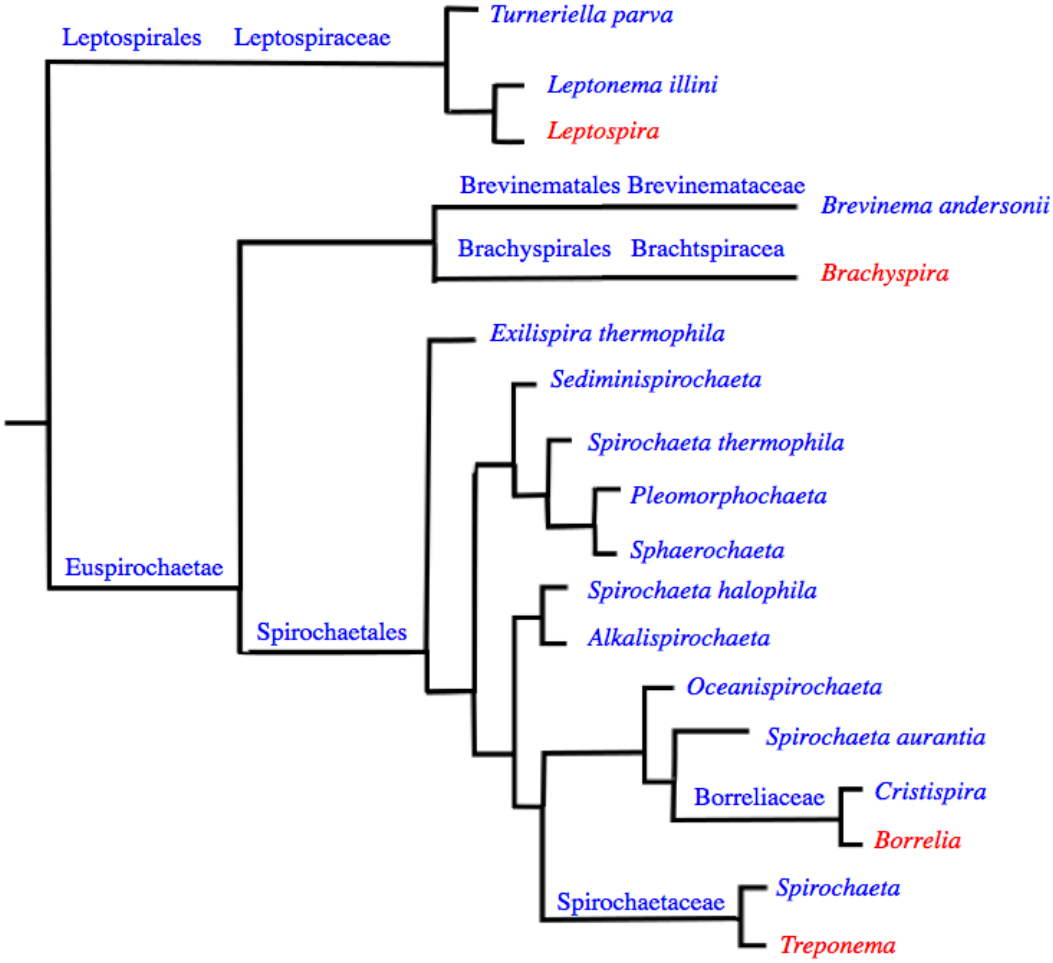


Figure 1.1: Phylogenetic tree based on 16S rRNA sequencing of species of the phylum Spirochaetes. Data obtained from the Living Tree Project (4). Genera containing pathogenic species are shown in red.

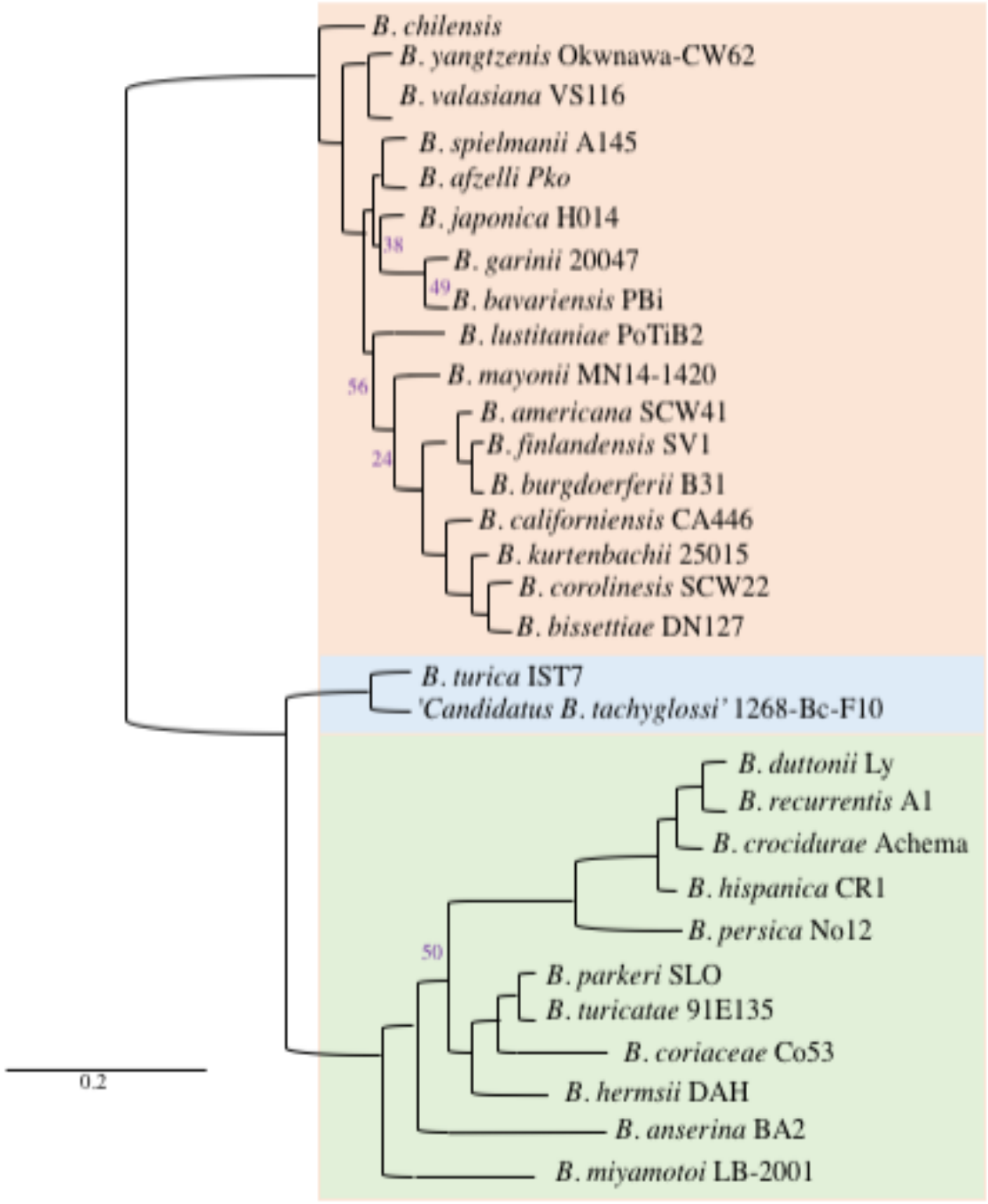


Figure 1.2.: Phylogenetic tree of 30 *Borrelia* species based on protein homolog data – Percentage of conserved proteins (POCP). Data used to create tree obtained from Margos *et. al* 2018 [3]. *Borrelia* strains associated with Lyme borreliosis/disease are highlighted in orange. Those associated with relapsing fever are highlighted in green. *B. turica* and *'Candidatus B. tachyglossi'* are from intermittent clades and are associated with reptile and echidna hosts.

N.B. All node support values are 100 except where indicated (in purple)

1.1.2. *Borrelia*

Borrelia is a genus of bacteria of the spirochaete phylum. Several species within the genus are responsible for the zoonotic diseases Lyme disease and relapsing fever. The major species causing Lyme disease are *Borrelia burgdorferi* sensu stricto (s.s.), *Borrelia afzelii* and *Borrelia garinii,* referred to collectively as *B. burgdorferi* sensu lato (s.l.). Causative species associated with Lyme disease are generally geographically distributed with *B. garinii* and *B. afzelli* responsible for the majority of cases in Europe, and *B. burgdorferi* s.s. responsible for the majority of cases in the United States. Other pathogenic species include *B. bavariensis, B. spielmanii and B. mayonii*. The s.l. group is continually expanding, with at least 20 recognised or proposed species (5). *B. burgdorferi* s.l. bacteria are spread to humans by tick vectors from a natural reservoir among small to medium mammals, reptiles and birds. Hard-bodied ticks of the genus *Ixodes* are the main vectors of Lyme disease. In Europe, mainly *Ixodes ricinus* and *I. persulcatus* hard ticks are implicated in the spread of Lyme disease. In the US, *I. scapularis* (deer tick) are the most common vector for Lyme disease (6).

Lyme disease was first formally clinically identified in 1975 following an outbreak of arthritis and juvenile arthritis in Old Lyme, Connecticut and surrounding areas. Initial studies by a group at the Yale School of Medicine, together with the Connecticut State Department of Health, found incidence of the now well-documented erythema migrans (EM) rash, clustering of symptoms coinciding with tick-feeding season and in some cases, known history of tick bite (7). While EM rash had previously been identified in Europe, it was suspected to be caused by localised-skin infection and had not yet been associated with incidence of arthritis. In the early 1980s, Willy

Burgdorferi identified spirochaetes in deer ticks and demonstrated this new bacterium, *Borrelia burgdorferi* is the causative agent of Lyme disease (8). In 1991, the disease was designated as a notifiable condition in the United States.

Generally, the spread and incidence of Lyme disease correlates to environments that facilitate the survival of *Ixodes* tick vectors. These includes temperate regions of the Northern Hemisphere. In Europe, distribution and density of *I. ricinus* and *I. persulcatus* generally correspond to the levels of Lyme disease seen in a particular region. The percentage of ticks that carry pathogenic species of *Borrelia* has been shown to vary in different geographical locations and this may be influenced by the species of tick that is dominant in that region, together with the small mammal population (9). Variation in percentage of ticks carrying *Borrelia* has also been shown to vary by season, with higher incidence seen during spring and summer months, coinciding with tick life cycle (10). Tick vectors have increased in density over time and have spread into higher latitudes and altitudes as the climate of these areas has changed (11, 12).

Lyme disease is the most commonly reported tick-borne disease in the United States and Europe. A total of 262,481 confirmed cases of Lyme disease were reported in the United States between 2007 and 2016 (13). In Europe, incidence varies widely by country, with 464 cases per 100,000 population seen per year in Sweden and 0.001 cases per 100,000 population seen per year in Italy (14). Lyme disease is not a statutorily notifiable disease in the United Kingdom. Since October 2010 under the Health Protection Regulations Act, every microbiology laboratory in England is required to notify all laboratory diagnoses of Lyme disease to Public Health England

(PHE) (15). In 2017, 1579 laboratory reports of Lyme disease were recorded in England and Wales (16). This figure is based on laboratory confirmed cases only and therefore the true incidence of Lyme disease, including unreported and misdiagnosed cases is likely to be significantly higher. In England and Wales, the incidence of laboratory-confirmed Lyme disease rose from 1.62 per 100,000 in 2013 to 1.95 cases per 100,000 in 2016 (17).

Due to the nature of infection via tick vectors, participation in outdoor activities is a risk factor in contracting the disease, particularly in high-risk areas including grassy and wooded areas in southern England and the Scottish Highlands; however, ticks in urban environments have also been shown to carry *B. burgdorferi* s.l. (8).

The *Borrelia* genus has at least two genospecies *B. turica* and *B. tachyglossi* that are not thought to be pathogenic to humans but are typically found in reptiles (18). A third group of the *Borrelia* genus contains over 29 species associated with relapsing fever. Tick-borne relapsing fever (TBRF) is a febrile disease, characterised by recurring episodes of fever and nonspecific symptoms, many of which are shared by Lyme disease including: myalgia, headache and arthralgia. TBRF is spread by exposure to the bite of an infected *Ornithodoros* tick. The tick species has a wide geographic spread and cases of TBRF are found throughout most of the world. It is endemic to several areas of North, Central and South America, the Mediterranean, Central Asia and much of Africa (19). TBRF has been associated with several small outbreaks (<62 cases) of infectious disease in the United States from the late 1960s onwards (20-24). The incidence of TBRF is difficult to estimate as many cases are unrecognised or misdiagnosed. TBRF is not endemic to the UK, and reported cases are assumed to be imported. Louse-borne relapsing fever (LBRF) is caused by *B. recurrentis* and is

spread by human lice vectors including *Pediculus humanus* and *P. capitis*. LBRF tends to occur in outbreaks and can cause severe symptoms with a fatality rate of 10-40% in untreated cases (25). LBRF is endemic in several East African countries including Ethiopia, Eritrea and Somalia. Cases in Europe are thought to be imported cases, often in refugees from East Africa (26).

1.1.3. *Leptospira*

Leptospira is a genus of the Spirochaetes phylum that contains several pathogenic species that together are responsible for causing the infectious human disease leptospirosis.

A total of 22 species of *Leptospira* have been identified, divided into three clades generally based on pathogenicity. At least 8 species of *Leptospira* are capable of being pathogenic in animals (27). Pathogenic species do not multiply in the environment, requiring an animal host (28). In 1915, *Leptospira* were first described as the causative agent of leptospirosis (Weil's disease) in humans (29).

Leptospirosis is found worldwide, but with incidence around 10 times greater in tropical and subtropical climates than that of more temporal climates (30). The bacteria have been shown to be capable of surviving for longer periods in warm and humid environments. Due to socioeconomic reasons, these areas are also less likely to have well developed surveillance systems for leptospirosis, together with conditions that facilitate greater transmission of *Leptospira* including inadequate sanitation and closer contact with animal species that are capable of carrying the bacteria. More cases occur during summer and autumn in temperate regions and during rainy seasons in warmer climates (31). Occupations that involve contact with carrier animals, or

with potentially contaminated water including farmers, veterinarians, rodent control works and sewer workers have a higher risk of contracting leptospirosis (32). Similarly, those participating in recreational activities that involve contact with water are at higher risk. The WHO has estimated an incidence of 873,000 cases worldwide annually, with over 40,000 deaths (33). Between 2007 and 2012, 301 cases of leptospirosis were confirmed in the UK, with around a quarter of these causes acquired abroad (34).

1.1.4. ***Treponema***

Treponema is a genus of spirochaete bacteria that contains the important human pathogen species, *Treponema pallidum*. Subspecies of *Treponema pallidum* are responsible for the diseases syphilis, bejel, pinta and yaws (35).

Syphilis is a chronic sexually transmitted infectious disease caused by the bacteria *Treponema pallidum* subspecies *pallidum (T.p. pallidum)*. While transmission of syphilis-causing *T.p. pallidum* is generally limited to close sexual contact or vertical transmission from mother to child in congenital syphilis, other pathogenic treponemes including *T. pallidum* subspecies *pertenue*, the causative agent of yaws, are transmitted by skin contact and development of infective lesions. Pinta, caused by *Treponema carateum*, and bejel, caused by *T. pallidum endemicum,* are chronic skin and tissue diseases.

T.p. pallidum was first identified in syphilitic sores (chancres) in 1905; however, evidence suggests syphilis or related diseases have been present for centuries and that an early European epidemic coincided with the return of Columbus from the New

World in 1493 (36). No animal reservoir exists for venereal syphilis, being a human only disease. There were an estimated 6 million new cases of syphilis worldwide in 2016. Incidence varies by country, with developing nations showing the largest burden of disease. The incidence of congenital syphilis in 2017 was 23.3 cases in 100,000 live births in the United States (37). Worldwide, between 700,000 and 1.6 million pregnancies are thought to be affected by syphilis, with the majority of perinatal deaths occurring in low-income countries. Incidence of congenital syphilis is very low in the UK, between 2010 and 2015, 17 cases were identified with no mortality (38). Incidence of venereal syphilis steadily decreased from the late 1940s due to the availability of antibiotic treatment; however, the number of cases has increased in recent years, with 7,149 cases diagnosed in 2017, up from the 2,874 cases diagnosed in 2008. A 20% rise in syphilis cases was seen between 2016 and 2017 alone (39) and a further rise of 5% to 7541 (40).

1.2 CELL STRUCTURE AND GENOME ORGANISATION

1.2.1. *B. burgdorferi* s.l.

B. burgdorferi s.l. is a helical shaped spirochaete, with a flexible cell wall that facilitates cell motility. The bacteria are typically 0.33 μm wide, and up to 25 μm long (41). As with all species within the spirochaete phylum, *Borrelia* have 7-11 internally located flagella between the inner and outer cell membranes. The flagella are structurally similar to flagella of other bacteria, comprising a basal-motor complex, hook and filament. Rotary motion action allows movement of the bacterium and

facilitates movement through host tissue (42). The flagellar filaments consist of repeating protein units encoded by the genes FlaA and FlaB (43).

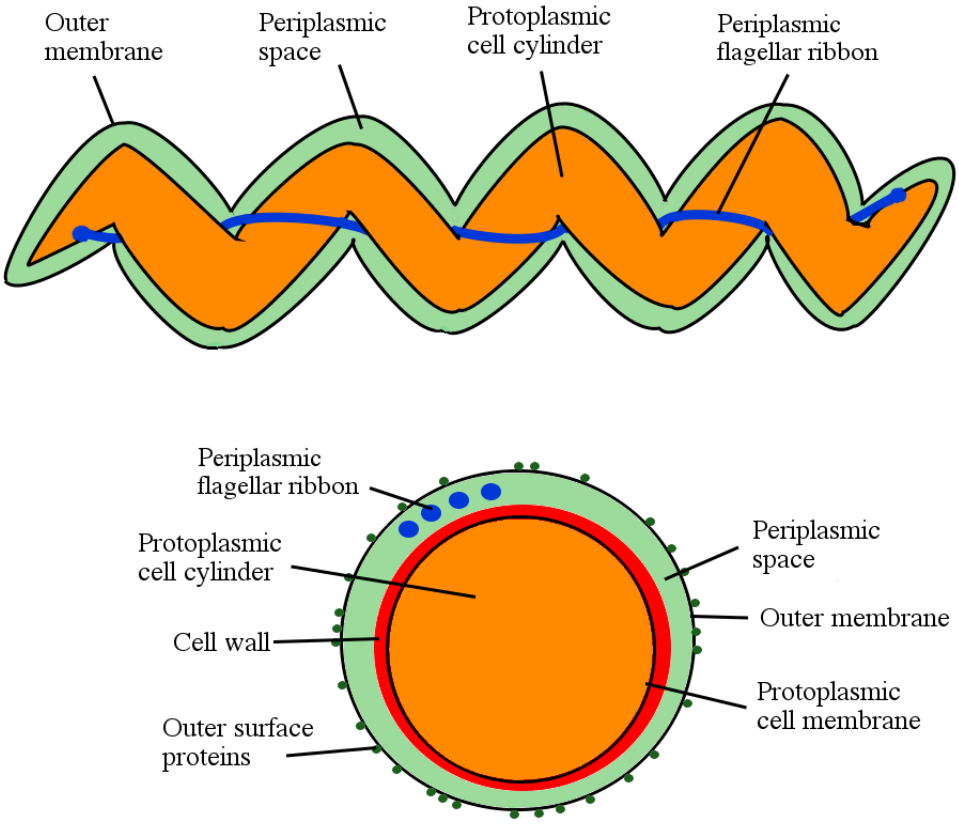


Figure 1.3: Schematic diagram of the structure and morphology of *Borrelia*. The helical shape of the bacteria is imparted by the presence of the flagellar ribbon. The cell has an inner and outer membrane as seen in the cross-section diagram. The periplasmic space between the membranes contains peptidoglycan (gram-negative) and the flagellar ribbon. Diagram author's own.

Borrelia express several outer surface proteins (OSPs) that have key roles in transmission between hosts. OspA is expressed on the surface of the bacterium whilst located within the tick midgut (44). Upon feeding, blood from the mammalian host enters the tick midgut, and expression of OspC is increased by the bacterium. *Borrelia* mutants lacking OspC have been shown to be competent to replicate and migrate to the salivary glands of the tick, but are not capable of infecting mice. OspC has been

shown to be indispensible for establishing infection in the mammalian host (45). Therefore, the reciprocal regulation of OspA and OspC facilitates pathogen transfer from the tick to its mammalian host and facilitates productive infection. The mechanism of action of OspA, OspC and other outer surface proteins is not fully understood in *Borrelia*, but their expression has repeatedly been shown to be vital for transmission between hosts and survival within diverse environments.

The genome of *B. burgdorferi* s.l. is one of the largest and most complex of any bacterium (46), likely a reflection of the wide range of genes that facilitate survival in tick and mammalian hosts. *Borrelia* have a ~950-kb linear chromosome with a number of circular and linear plasmids that can vary between species and subspecies. *B. burgdorferi* s.s. contains over 20 linear and circular plasmids (47). The majority of genes associated with encoding outer membrane lipoproteins are located on the plasmids (48).

Despite their large genome, *Borrelia* lack the capacity to synthesize amino acids, fatty acids, nucleotides and several enzyme cofactors. Therefore, *Borrelia* are reliant on obtaining these from the host, and the bacteria have a wide range of genes encoding transporters and binding proteins for carbohydrates, peptides and amino acids (49, 50). *Borrelia* remains the only known organism that does not require iron for growth (51). During bacterial infection, as part of the innate immune response, host cells typically increase production of lactoferrin to limit availability of iron together with several other mechanisms, including the production of iron sequesters such as lipocalin 2 (52). Lacking requirement for iron, *Borrelia* have greater immunity to

these host defences. Previous studies have shown that for several functional proteins, *Borrelia* have adapted to use magnesium as a cofactor rather than iron (53).

1.2.2. *Leptospira*

As spirochaetes, the *Leptospira* share several key cell structures with *Borrelia* including a long, helically coiled shape; however, leptospires differ from other spirochaetes as they have single flagellum rather than a bundle (54). The outer membrane of the *Leptospira* contains a wide range of lipoproteins and transmembrane OSPs that are associated with pathogenicity. OSPs have been shown to be particularly important in pathogenic species of *Leptospira* for attaching to host extracellular matrix as well as to immune mediators including the complement protein Factor H (55).

The genome of *Leptospira* consists of two chromosomes that vary in size depending on the particular species. The genome of *Leptospira interrogans* consists of a 4.33 Mb circular chromosome and a 350 kb small chromosome (56). The genome of pathogenic species has a large proportion dedicated to pathogenic related genes. The *L. interrogans* genome contains a wide range of pathogen-specific genes together with a host of housekeeping genes that facilitate survival of the bacteria in diverse environments from stagnant water to mammalian host (57).

1.2.3. *T.p. pallidum*

T.p. pallidum share the helical structure and internal flagella that facilitate their transmission into human hosts via a corkscrew motion through mucous membranes or breaks in the skin. Unlike *Borrelia* and *Leptspira*, however, *T.p. pallidum* has an outer

membrane that has very few integral membrane proteins. Due to the fragile, sparse nature of membrane proteins in syphilis, few have been elucidated and their influence on infection is not fully defined (58).

In contrast to the large, complex genome of *Borrelia, T.p pallidum* has one of the smallest bacterial genomes at 1.14 Mb and does not synthesise lipids, nucleotides and the majority of amino acids and are therefore highly dependent on the host for nutrients (59). The very small genome in relation to *Borrelia* and the *Leptospira* concurs with its inability to survive outside its obligate human host. The bacteria are highly adapted to evade recognition by the host's innate and adaptive immune responses to facilitate colonisation and chronic infection. *T. pallidum* utilises most of the essential substrates from the host environment using a large number of specific transporters (60).

1.2.4. Genome diversity amongst the spirochaetes

Whilst showing great diversity in pathogenicity, transmission and niche, the spirochaetes share several structural and genomic features. A comparison of *L. interrogans, B. burgdorferi* s.l. and *T.p. pallidum* showed that 41% of the genes found in *B. burgdorferi* s.l. were also found in *L. interrogans,* and that 57% of the genes in *T.p. pallidum* are found in the *L. interrogans* genome. 362 genes were found to be shared by all three spirochaetes (61).

1.3 TRANSMISSION AND PREVENTION

1.3.1. *Ixodes* tick life cycle and the transmission of *B. burgdorferi* s.l.

In Europe, the *Ixodes ricinus* tick is the most common vector for transmission of *Borrelia spp*.. The bacteria can be transmitted to between small and medium mammalian hosts and to humans when the tick vector attaches and takes a blood meal. The tick is found across Europe, commonly in areas with abundance of appropriate reservoir hosts. In particular, remote forested areas and heathland have high incidence of ticks (62); however, they are also present in some urban and suburban habitats and have previously been sampled in South London parks (63). While the risk of transmission from ticks is greatest amongst those who work or spend significant amount of time in rural locations, including farmers and hill walkers, there is still measurable risk of tick-bite throughout the UK, wherever *Ixodes* ticks and reservoir species exist.

Ixodes ticks are widely distributed throughout the Northern Hemisphere with their range continually expanding due to reforestation and increasing temperatures providing habitable conditions for new populations of host mammals (64). In the UK, a 2008 study found that *Ixodes ricinus* distribution had increased by 17% since 2005, with increases in tick numbers coinciding with local increases in deer population (65, 66). Public Health England's passive Tick Surveillance Scheme found *Ixodes ricinus* to be the most frequently recorded endemic tick species in the UK at 59.2% of all recordings (67). *Ixodes ricinus* has a three-host life cycle with four life stages: egg, larva, nymph and adult **(figure 1.4).** In order to produce eggs and moult, female ticks require a sufficient blood meal and can feed on a wide range of hosts including small

rodents during earlier stages of their life cycle often progressing to larger mammals such as roe deer at adult stages of development. Large mammals are essential for the survival of most tick populations, providing sufficient feed for large numbers of adult ticks and facilitate distribution of ticks and eggs into new areas; however, they are not competent reservoirs are spirochaetes. The main reservoirs for *Borrelia* in both Europe and the United States are small mammals including mice, voles and some species of bird (68).

The ecology and biodiversity of an area, particularly the population size of competent *Borrelia* reservoir species influences the positivity rate in ticks and ultimately the risk of contracting Lyme disease The percentage of ticks that carry the *Borrelia* pathogen varies widely depending on geographical location and species of tick. A large scale study in the United States involved the screening of 16,080 ticks submitted by the public for several species of *B. burgdorferi* s.l. found a prevalence of 19.5% in adult ticks, 11.0% in nymphs and 5.1% in larvae (69). In the UK an assessment to the suitability of urban greenspace for ticks such as urban woodlands showed a mean prevalence of *Borrelia*-infected *I. ricinus* ticks at 18% (70). The majority of humans are infected through the bites of nymphs as larger adult ticks are more likely to be identified and removed before transmission. A 2018 study found an approximate 10% transmission rate by 48 hours of nymph tick attachment, rising to 70% by 72 hours (71). In North America, *Ixodes scapularis* and *I. pacificus* are the most common vector for *Borrelia* and are also capable of transmitting the causative agents of anaplasmosis and babesiosis. *I. persulcatus* represent the major vector for *Borrelia* in Asia (72). **Figure 1.4** shows a diagram of the lifecycle of *Ixodes* ticks.

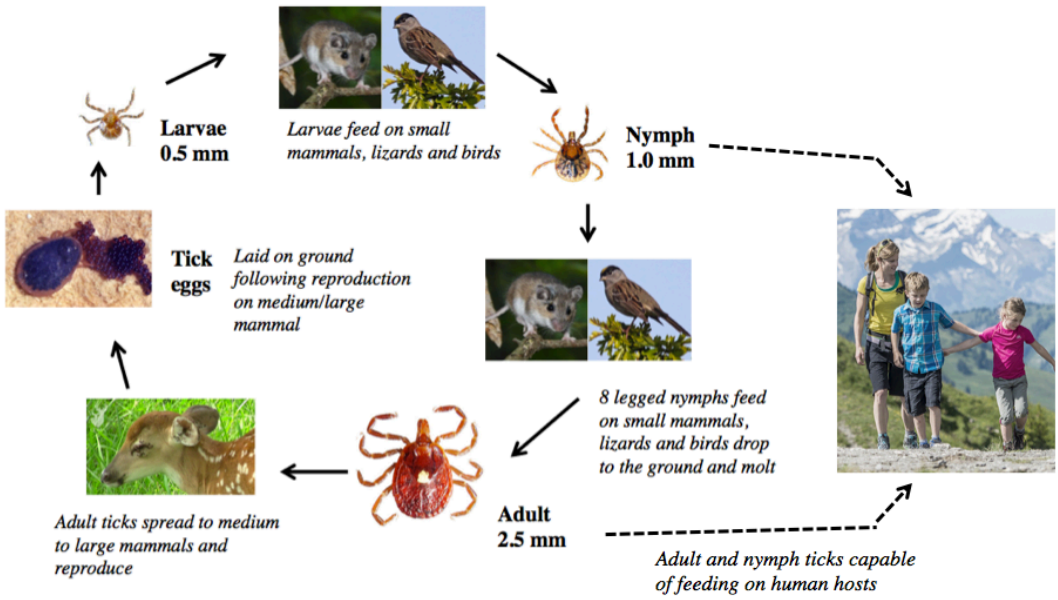


Figure 1.4: Lifecycle of the *Ixodes* tick vectors of *B. burgdorferi*. Small mammals are confirmed resorvoirs for pathogenic species of *Borrelia* and pass on the spirochaete to tick vectors during nymphal feeding

Preventative methods such as encouraging the use of insect repellent on exposed skin, wearing protective clothing and guidelines on the prompt and safe removal of ticks form the basis of public health protection strategies (73).

There is no vaccine currently in use for Lyme disease. LYMErix™, a recombinant Lyme vaccine, was approved for use by the FDA in 1998. The vaccine was in production for until 2001 when the manufacturer voluntarily withdrew the vaccine from market following reports of adverse reactions and widespread media coverage (74). 1.4 million Lyme vaccine doses had been administered by this point. Studies have not substantiated many of the safety concerns associated with the vaccine (75, 76).

1.3.2. Transmission of *Leptospira*

In the majority of leptospirosis cases, transmission of the pathogen is through abrasions or cuts in the skin or via the conjunctiva of the eye (77). Pathogenic strains of *Leptospira* are capable of surviving in stagnant water for many days to weeks (78) and water supplies can be contaminated by the urine of an infected mammalian host. Spending prolonged periods of time in contaminated water is the major risk factor for contracting leptospirosis (79). Inhalation of aerosols containing the spirochaetes can also lead to infection through the respiratory tract mucous membranes (77). Lack of sanitation, poor housing and inadequate water treatment facilities contribute to the endemic levels of leptospirosis seen in tropical and subtropical regions, together with weather conditions that facilitate survival of the spirochaetes. The bacteria is carried by several animal species, with rodents representing the most important primary hosts in the majority of human infections (80). Prevention of leptospirosis is based on spreading awareness of transmission risks; particularly those associated with entering potentially contaminated water supplies. Local infrastructure, including water supply, rodent prevention and housing standards are important for the prevention of leptospirosis (77).

1.3.3. Transmission of *T.p pallidum*

Unlike *Borrelia* and *Leptospira* the causative agent of syphilis, *T.p. pallidum* is an obligate human pathogen with no known animal reservoir or vector. Syphilis is transmitted primarily by direct sexual contact with the infectious lesion of another person. Cases of acquisition through blood products and organ donation have been reported (81, 82), but these are rare incidences. Congenital syphilis is spread *in utero* or during birth from an infected mother (83). Transmission of congenital syphilis can

generally be prevented successfully by treatment of infected mothers using antibiotics prior to childbirth (84). For venereal syphilis, prevention strategies rely on safe sex practices, including the use of appropriate prophylactics, Sexually Transmitted Infection (STI) screening availability and public awareness of STI risks. The availability of pre-exposure prophylaxis medication (PrEP) for HIV, has resulted in a decrease in condom use particularly in men who have sex with men. This may contribute to the large increase in syphilis cases reported within this group over recent years (85).

1.4 DISEASE PATHOGENESIS

1.4.1. Lyme disease

In human hosts, the bite of an infected *Ixodes* tick is required for the introduction of *B. burgdorferi* s.l. through healthy skin. Colonisation and persistence in the dermal layer allows adaptation of the pathogen through altered gene expression facilitated by the movement from the environment of the tick midgut to the human host (86, 87). Broadly, Lyme disease can be split into three stages; early localised Lyme disease, early disseminated Lyme disease and late disseminated Lyme disease. The infecting *Borrelia* genospecies influences the likelihood of development of certain sequelae during dissemination due to differing tissue tropism. **Figure 1.5.** shows the progression of Lyme disease through these stages. It should be noted that Lyme disease is a complex condition and several of the symptoms described have at least some element of host autoimmunity; therefore, a wide range of symptoms have been attributed to Lyme disease, several of which are shared with other infectious and non-

infectious disease. The pathogen does not produce any known toxins or proteases that could be directly responsible for tissue damage seen during colonisation (88). Here, well-documented conditions associated with *Borrelia* infection are given along with estimations of their incidence in Lyme disease patients. In reality, Lyme disease patients are likely to present with a complex set of symptoms, particularly in the later stages of disease.

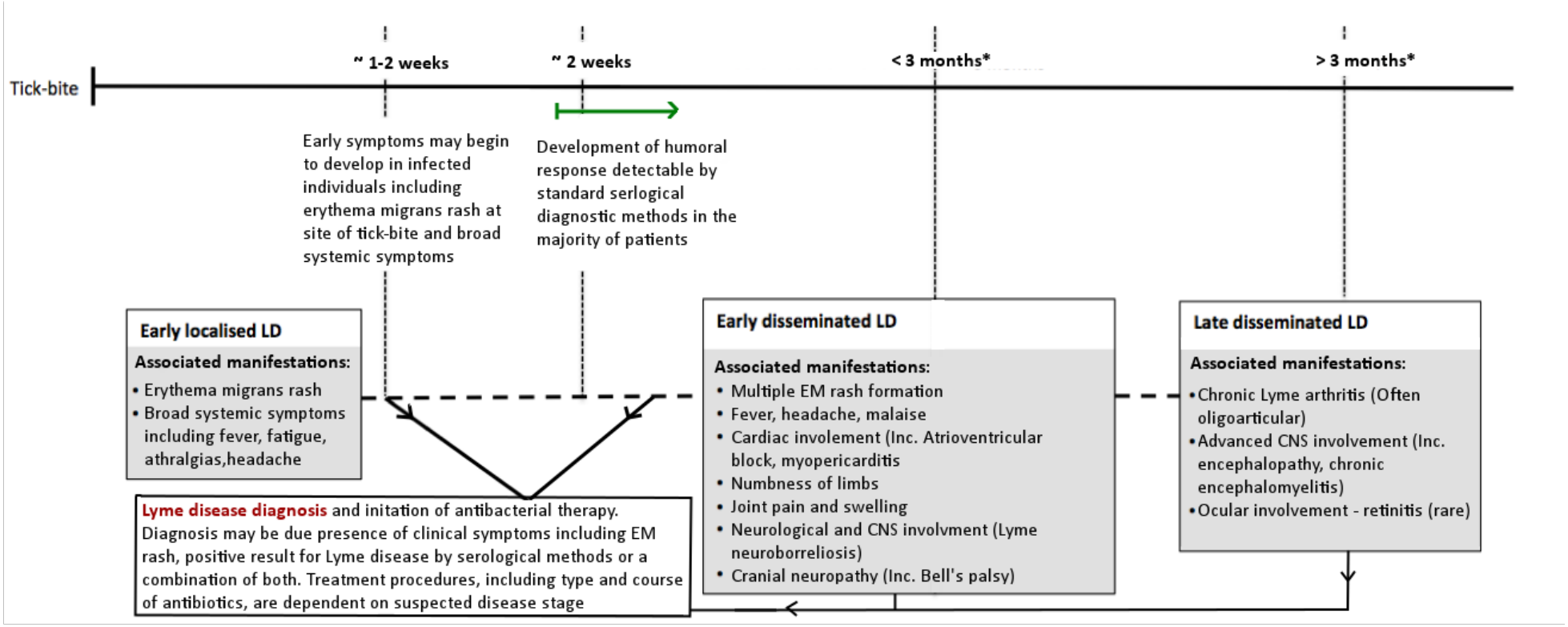


Figure 1.5: Summary diagram of the progression of Lyme disease from time of tick bite to complications associated with late stage of untreated disease. *The time frames provided give a general guide on when different stages occur, based on available literature. Lyme disease patients can develop specific manifestations at atypical time points. Course of disease is influenced by a combination of factors including the infecting *Borrelia* genospecies, host immunity and treatment interventions.

1.4.1.1. Early localised Lyme disease

The presence of *Borrelia* bacteria in the dermal layer stimulates host proinflammatory immune responses that cause a radiating rash around the site of tick-bite that subsequently clears from inside, causing the characteristic erythema migrans (EM) rash. While successful colonisation of the dermal layer is requisite for Lyme disease and progression to later stages of disease, EM rash is not seen in all patients. The incidence of EM rash in Lyme disease cases is thought to be around 70%, although estimates vary widely depending on the study (89). EM rash can vary greatly in shape and may include vesicular or necrotic areas towards the centre, further complicating its recognition. EM alone is usually asymptomatic, but it may be painful and accompanied by other symptoms including fever, generally feeling unwell, headache and myalgia. If diagnosed during the early, localised stage, treatment in the UK should involve a 3 week course of oral antibiotics such as doxycycline (90). Antibiotic treatment has a high success rate when administered during this stage (91). Lyme disease may progress to a disseminated form when early, acute cases are missed due to asymptomatic presentation, or due to misdiagnosis.

1.4.1.2. Early disseminated Lyme disease

In the majority of cases, the innate immune system is capable of controlling bacterial burden in the dermis but is often unable to clear infection allowing for the dissemination of *Borrelia* and progress of disease. Dissemination to secondary colonisation sites requires continual and wide-ranging adaptation to overcome environmental changes and to allow successful evasion of the host immune response. can spread through non-haematogenous routes including by direct spread though tissue and the lymphatic system while dissemination through the circulatory system

allows colonisation of more distant sites (92) The most common manifestation of early, disseminated Lyme disease seen in the United States is the presence of multiple EM rashes, distinct from the original site of tick bite. The presence of multiple EM rashes is significantly more common in North America than in Europe (93). These secondary lesions appear around 2 weeks after the development of primary EM (94). Other symptoms associated with Lyme disease include Bell's palsy and other cranial nerve palsies and meningitis. Neurological symptoms are more common in European cases of Lyme disease (95). Rarely, early stages of infectious carditis can also occur (96). Systemic symptoms during dissemination are again wide-ranging and non-specific to Lyme disease including fever, myalgia, headache and fatigue. Treatment of Lyme disease that has progressed to the early dissemination stage often depends on the symptoms observed. The presence of multiple EM rashes, cranial palsy and early carditis are often still treated by oral antibiotics. In more severe cases of palsy or carditis, intravenous courses of antibiotics may be administered (97). UK NICE guidance for treatment of early disseminated stages of Lyme not affecting the Central Nervous System (CNS) involves oral doxycycline for 21 days (90).

1.4.1.3. Late disseminated Lyme disease

Several factors influence the progression of disease from this early, localised disease to dissemination and the development of sequelae associated with later stages of Lyme disease. These include whether antibiotic treatment has been administered, together with host immunological factors. Without medical intervention, early, disseminated Lyme can progress to late stage disseminated Lyme disease. This stage of disease is often characterised by an increase in severity of previous systemic symptoms including headaches and fever, but is also associated with the development

of chronic conditions including Lyme arthritis, neuroborreliosis and more rarely, acrodermatitis chronicum atrophicans (ACA) or carditis. Encephalomyelitis and polyneuropathy are rare late stage manifestations (98). In the United States, Lyme arthritis is the most commonly recorded symptom of late stage infection, usually occurring several months following tick bite. Lyme arthritis can be intermittent or persistent depending on severity. Without prior antibiotic treatment, it is estimated 60% of patients would develop Lyme arthritis (99), although several factors including host immunity and infecting genospecies of *Borrelia* are likely to influence the incidence of arthritis. The arthritis is usually monoarticular, affecting the large joints, but other manifestations of arthritis are not uncommon. In Europe, the condition acrodermatitis chronica atrophicans (ACA), a chronic dermatitis associated with EM rash is a common feature of late stage disseminated Lyme disease Prevalence of ACA in all European patients with Lyme disease is estimated to be 1-10%, although this figure is likely to vary widely depending on geographical location (100, 101). Neurological complications associated with Lyme disease, collectively known as neuroborreliosis, can occur during late stages of disease. Neuroborreliosis is more common in European patients, and often includes painful meningoradiculitis. More severe cranial nerve symptoms occur in around half of patients with neuroborreliosis, often with unilateral or bilateral facial nerve involvement (102). Symptoms of CNS involvement, including encephalitis and myelitis are rare in patients with early stage neuroborreliosis, but risk increases in absence of medical intervention (103).

Treatment of the majority of late stage manifestations of Lyme disease including Lyme arthritis, early neuroborreliosis and ACA is usually by a 3-4 week of oral antibiotics. In more severe cases, or in the case of CNS involvement, intravenous

antibiotics may be prescribed. The benefit of long-term antibiotic therapy in the treatment of late stage Lyme disease has been widely discussed, but current evidence does not support any benefit in long-term treatment outweighing the well-documented risks associated with prolonged courses of antibiotics (104).

1.4.1.4. Post-treatment Lyme disease syndrome

The term post-treatment Lyme disease syndrome (PTLDS) has been used to describe continuing or relapsing symptoms in patients with previously treated Lyme disease. Some studies have shown up to 10% of patients with previously documented Lyme disease continue to report musculoskeletal or cognitive dysfunction after recommended antibiotic treatment (105). Most of these issues resolve with time, but in some cases, symptoms attributed to PTLDS have been reported several years after tick bite (106). PTLDS has not been fully defined, and the term is often used interchangeably with Chronic Lyme disease (CLD). The Centres for Disease Control and Prevention (CDC) in the United States describe PTLDS as: "Symptoms of pain, fatigue, or difficulty thinking that lasts for more than 6 months after a patient finishes treatment" (107). Other national guidelines on diagnosis and management of Lyme disease, including those of the National Institute for Health and Care Excellence (NICE), have avoided using either term and have concentrated on providing advice on treatment and diagnosis based on available evidence (83). Regardless of definition, the underlying biological mechanisms that result in some patients experiencing persisting symptoms following Lyme disease are not known. Many persistent symptoms reported in PTLDS share characteristics of autoimmune disease, potentially triggered by Borrelial infection. Other groups believe that ongoing symptoms can be attributed

to a persistent, but difficult to detect, infection. Potential mechanisms of persistence in the host by *Borrelia* are discussed in the next section (1.4.1.5.).

It can be argued that documented Lyme disease diagnosis and/or positive serology are requisite features of patients with PTLDS; however; misdiagnoses, false negative serology, and prescription of antibiotic treatment without serological confirmation of causative agent do occur in Lyme disease cases. Significant research challenges need to be met to understand how Lyme disease may be attributable to the development of any long-term symptoms.

1.4.1.5. Chronic Lyme disease

Chronic Lyme disease (CLD) can be broadly defined as the presence of persistent long-term symptoms that are associated with Lyme disease. There is no agreement on the exact criteria that constitutes a CLD patient, or an accepted clinical definition. The term has grown to cover a broad range of symptoms, some of which are associated with Lyme disease, including peripheral neuropathy and arthralgias (108); however, presence of more general symptoms including fatigue, cognitive dysfunction and headaches are too broad to rule out numerous other diseases. Without a definition for CLD, studies published using populations with putative CLD as a cohort seem to hold little value. The broad range of symptoms documented in patients with Lyme disease, together with the fact that certain manifestations of the disease including Lyme arthritis are chronic in nature, has likely led to cases of putative PTLDS or long-term symptoms of disease being missed from research, while patients suffering from other maladies are misdiagnosed as having CLD. Most criteria given of CLD, including those of the International Lyme and Associated Diseases Society (ILADS) (109), do

not include prior positive serology for Lyme disease, history of tick-bite or presence of certain distinguishing features including EM rash. The diagnosis of Lyme disease by conventional serology and by other more novel testing, including methods often associated with CLD diagnosis, are discussed in section 1.6.

The underlying pathophysiological aspects of persistent symptoms attributed to Lyme disease are not well understood and remain a controversial subject. In other infectious diseases, some prior studies have shown higher risks of developing certain autoimmune diseases, including lupus in patients with prior Epstein-Barr virus (EBV) infection (110). The role of autoimmunity in Lyme disease is currently not well understood. ILADS and other groups attribute PTLDS/CLD to persistent infection with *Borrelia* s.l.. Several theories have been proposed to explain how the bacteria could persist whilst avoiding the host immune system, antibiotic treatment and detection by diagnostic methods. Some groups propose that *Borrelia* s.l. are capable of becoming an intracellular pathogen, penetrating host cells (111). *Borrelia* s.l. have previously been shown to predominantly occupy the extracellular matrix (112), and have been shown to disrupt pathways associated with host cell cytoskeletal rearrangement, including of collagen, in murine models (113, 114). While disruption of host cell structural mechanisms can be associated with intracellular pathogens including *Listeria monocytogenes* (115), *Rickettsia rickettsii (116)* and *Shigella (117)*, the presence of similar mechanisms in *Borrelia* s.l. may only promote establishment within the extracellular matrix. *Borrelia* s.l. have been shown to be capable of invading human primary fibroblasts and endothelial cells following long co-culture in an *ex vivo* assay (118); however, sound peer-reviewed evidence demonstrating internalisation of *Borrelia* s.l. by non-immune system cells during infection *in vivo* is

lacking. *Borrelia* s.l. may be capable of forced cell internalisation within the human host; however, the influence, if any, of this ability on persistence in the host or on disease progression is not known. Proponents of the intracellular persistence theory argue that *Borrelia* s.l. may evade being killed by antibiotics by hiding within host cells. In contradiction, doxycycline and other commonly prescribed antibiotics for Lyme disease are also currently recommended for the first-line treatment of *Rickettsia* (119) and other facultative intracellular bacterial pathogen (120) infections.

There are numerous study design issues surrounding current publications promoting the use of extended antibiotic therapy for treatment of PTLDS or CLD, including a study featuring intravenous antibiotic delivery for up to 52 weeks (121). No conclusive evidence has shown that long-term use leads to substantial improvement in patient outcomes (91). The potential dangers to the patient surrounding long-term antibiotic use for Lyme disease are well-documented (122-125), and such therapy may drive development of antibiotic resistance (126). Numerous clinical guidelines from agencies and groups representing countries with endemic Lyme disease recommend against the use of prolonged antibiotic treatment (126-137).

1.4.2. Leptospirosis

Upon human contact with contaminated soil or water, the spirochaetes can enter the body through breaches in the skin, or through mucous membranes, then enter the bloodstream. A risk factor for contracting leptospirosis in humans is presence of leg wounds that allow transfer of the pathogen from a contaminated water source, often when performing tasks such as clothes washing (138). Upon entering the bloodstream,

the leptospires attach to endothelial cells of the blood vessel walls and subsequently move between cell layers facilitated by the corkscrew motion of the pathogen's internal flagella. *L interrogans* is capable of attaching to a number of human cell types, but importantly has high affinity for kidney epithelial cells (139). Upon entering the host, Microbial Pathogen-Associated Molecular Patterns (PAMPs) on the bacterial surface can be recognised by Pattern Recognition Receptors (PRRs) of the innate immune system, causing activation of inflammatory responses. The upstream regulator proteins NF-kappa-B and activator protein 1 (AP-1) have been shown to be important in activation of immunological cascades in leptospirosis (140). Leptospirosis patients show increased expression of pro-inflammatory cytokines including IL-6, IL-12, IFNs, tumour necrosis factors (TNFs) that recruit cells of the immune system to the site of bacterial attachment. Severe forms of leptospirosis are often marked by dysregulation of immune and inflammatory response, often called a cytokine storm, that cause persistent inflammation and production of pro- and anti-inflammatory cytokines (141).

Clinical manifestations of infection are highly variable in humans and in 90% of cases are mild or asymptomatic and resolve spontaneously (142). In the 10% of cases that show persistent infection, symptoms can be very severe and may affect multiple organs. Systemic symptoms of infection include febrile illness of sudden onset, headache, myalgia and, rarely, a skin rash. In such cases of icteric leptospirosis with jaundice, the clinical course is often rapidly progressive. Complications of severe leptospirosis are wide ranging, a reflection of the multisystem nature of infection. An important manifestation of leptospirosis is kidney damage and acute kidney injury (AKI), which occurs in up to 40% of icteric cases (143). AKI is characterised by

tubule dysfunction, hypokalemia and cell necrosis. Activation of inflammatory processes by recognition of bacteria by toll-like receptors (TLRs) elicits tubular inflammation and damage and can also lead to the accumulation of extracellular matrix in the kidney (144). Ocular manifestations are seen in a significant number of cases, with incidence reaching up to 40% in certain high-risk groups (145). Conjunctival suffusion is not commonly found in other febrile illnesses and its presence, along with sceleral icterus (eleveated serum bilirubin causing yellowing of the eye) is considered pathognomic of severe leptospirosis (145). Jaundice is also commonly reported in patients with severe leptospirosis, likely to be caused by the host inflammatory response. Necrosis of muscle fibres, particularly in skeletal muscles is also observed in many cases and may account for the intense myalgia reported by some patients (146). Severe leptospirosis can also affect the brain and spinal cord with potential development of meningitis and/or encephalitis (147).

1.4.3. Syphilis

Venereal syphilis can present in four stages: primary, secondary, latent and tertiary. Primary infection, most often acquired by sexual contact with the infectious lesions of a sexual partner is marked by the development of chancre at the point of contact after 10-90 days (60). The helical nature and flagella of *T.p. pallidum* facilitates access of the bacteria into the host through breaches in squamous or columnar epithelium. In heterosexual men, chancres most often occur on the penis, but can occur in other sites including the rectum and oral cavity. In women, primary lesions usually occur on the labia or cervix. A significant proportion of lesions are painless, leading to delayed diagnosis in a number of cases (148). Secondary syphilis occurs a number of weeks following primary infection and is marked by development of symptoms including a

non-itchy rash on the skin, condyloma latum lesions on mucous membranes and systemic symptoms such as fever, sore throat and general malaise (149). Latent syphilis is described as the period of time when no visible signs or symptoms of syphilis occur, but the pathogen is still living in the host in a dormant state. Latent syphilis occurs as early as 1 year following primary infection and, depending on the definition, can last up to 2 years. Depending on the stage of latency, subjects can still be contagious during this time (150). Tertiary syphilis relates to long-term, end organ complications associated with syphilis infection, occurring several years after primary infection. Around 35% of patients with latent syphilis will go on to develop tertiary syphilis. Complications involving the CNS (neurosyphilis), the cardiovascular system (cardiovascular syphilis) and non-cancerous growths (gummas) are the main observations seen with tertiary syphilis. Penicillin injections remain as the most common and effective method for treatment of all stages of syphilis (151).

1.5. DIAGNOSTICS FOR LEPTOSPIROSIS AND SYPHILIS

1.5.1. Leptospirosis

The most common method for the diagnosis of leptospirosis is serology. Historically, the Microscopic Agglutination Test (MAT) was considered the gold standard of leptospirosis diagnosis as it is highly sensitive and allows for the detection of leptospirosis group (serovar)-specific antibodies (152). The test involves the use of live leptospires, ideally recent isolates representing serovars circulating in the area where the disease was contracted. Patient serum is mixed

with the leptospires and if antibodies against leptospiral antigens are present in the serum, the bacterial cells will become agglutinated. The test is less sensitive in early phases of disease and its complicated and labour intensive process can lead to the introduction of human error into testing protocols. It is therefore best suited to reference microbiology applications.

The most widely used serological test for leptospirosis diagnosis is an ELISA based IgM assay. The test detects the presence of Leptospira specific IgM antibodies in human sera by a sandwich-ELISA method and is capable of detecting infection caused by a number of *L. interrogans* serovars including *hardjo, Pomona, copenhageni, australis* and *tarrasovi* (153). Disadvantages of the IgM assay include low sensitivity in early stages of infection and persistence of IgM antibodies potentially leading to false positives in patients with resolved leptospirosis (154).

In recent years, several real-time polymerase chain reaction assays have been described. These can confirm the diagnosis in the early phase of the disease before antibody titers are at detectable levels. In the UK, IgM ELISA is used in conjunction with a UKAS accredited real time PCR assay for the detection of Leptospira DNA and for differentiation between pathogenic and non-pathogenic species or serovars. The test involves a primary PCR that targets the bacterial outer membrane lipoprotein LipL32, followed by a second triplex reaction that targets a conserved region within the 16S rRNA gene (rrn) in Leptospira. Labelled probes within this triplex reaction closely correlate to pathogenic, intermediate and non-pathogenic leptospires (155). Samples that test positive

for leptospirosis by IgM ELISA or PCR in the UK are forwarded for MAT testing for reference purposes and speciation.

Culture of Leptospira is a difficult, laborious process that requires the use of expensive medium and may take up to several months; therefore culture of Leptospira is only undertaken retrospectively, often in research settings and is not recommended for routine diagnosis (156). Dark field microscopy is also capable of visualising leptospirosis, but the sample must be taken from the site of current infection and it lacks sensitivity and specificity. High bacteraemia (>10 leptospires/mL) is required for successful visualisation of Leptospira under dark field microscopy and the process is labour intensive and can lead to human error resulting in false positives (157).

1.5.2. Syphilis

In most countries, blood serological testing remains the diagnostic standard for syphilis. The standard processes involves a nontreponemal test (NTT) followed by a treponemal test (TT) (158). Detection of at least one NTT antigen and one TT antigen is sufficient to confirm syphilis. NTTs, such as the rapid plasma reagin test (RPR) are used to monitor the state of infection by measuring levels of IgG and IgM antibodies produced in response to material released from damaged host cells, primarily cardiolipin. As the test measures IgG and IgM, the test is relatively effective even in early infection, with 75% sensitivity seen during primary syphilis (158). NTTs can also be used to monitor progression of disease and response to treatment. TT tests are specific to *T. pallidum* infection and detect the presence of pathogen-specific antibodies. The commonly used *T. pallidum* hemagglutination assay (THPA) and *T.*

pallidum particle agglutination assay (TPPA) are indirect agglutination tests in which bacterial surface antigens are coated onto red blood cells or gelatin particles respectively. Patient serum is mixed with these substrates and, if sufficient specific host antibodies to *T. pallidum* are present, agglutination occurs. The TTPA test is one of the first tests to become reactive in primary syphilis and is often used as an initial test together with specific IgM (159). Western blot methods can be used for detection of specific antibodies for both IgM and IgG with many tests now using immunoblots of recombinant antigens derived from *T. pallidum* (160).

No centralised system for diagnostic testing exists for syphilis in the UK, with testing often conducted locally or on site following presentation with suspected syphilis. In suspected early primary syphilis, PHE guidelines recommend sampling of syphilitic lesions (chancres) for treponemal PCR together with dark ground microscopy where available (161). For serological testing, the stage of disease may influence the most suitable test. Generally, it is recommended that an initial test for treponemal specific antibodies is undertaken (EIA or CLIA detecting IgG and IgM). In the case of negative results, retesting may be appropriate particularly if tested within 2-4 weeks of infection, as the patient may not have seroconverted. In the case of positive EIA or CLIA, a second confirmatory test is recommended and may involve TPPA, TPHA or TPLA methods. As with other serology based diagnostic methods, laboratory test results must be considered together with the clinical and geographical background of the patient.

1.6. DIAGNOSTICS FOR LYME DISEASE

1.6.1. PCR and culture for diagnosis of Lyme disease

Quantitative real-time PCR tests have been extensively developed in recent years for the routine diagnosis of bacterial diseases. PCR based tests detect the presence of specific bacterial DNA directly and therefore do not require the development of an antibody response. In Lyme disease, the blood-borne phase of dissemination is relatively brief and the number of spirochetes in infected tissues remains low even in late stages of disease. For this reason, the efficacy of PCR-based tests for Lyme disease diagnosis is highly dependent on the type of biological sample collected. The sensitivity of PCR for detection of spirochete DNA in EM lesion samples, where the bacteria are actively replicating, is relatively high. In a study using skin biopsy samples from patients with EM, DNA was detected in 71% of samples (162). Several other studies in Europe and the United States have shown similar levels of sensitivity in skin biopsy based PCR tests (162-164). While the sensitivity of PCR for EM skin biopsy samples is comparatively high, particularly in cases of acrodermatitis chronica atrophicans, performing such biopsies is invasive, not routine and is unlikely to be applicable to clinical diagnosis on a wide scale. As most studies only show high levels of sensitivity for PCR in patients with EM, it is unlikely that PCR will have value in the diagnosis of patients with atypical symptoms. A positive result by PCR confirms detection of pathogen DNA but does not distinguish between viable or dead bacteria.

The direct culture of viable spirochetes from clinical specimens remains the gold standard for diagnosis of Lyme disease. In routine diagnosis, however, culture is rarely utilised. Culture of *B. burgdorferi* requires expensive Barbour-Stoenner-Kelly

medium and, due to the slow replication rate of the bacteria, is a time-consuming process. Recovery of spirochetes from 2-4mm biopsy samples of EM lesions can be achieved in at least 40% of untreated patients, with some studies reporting a success rate of above 86% (165). Recovery of bacteria from blood, serum and plasma is unusual, with success rates generally below 5%. While recent advances in culture techniques have improved the sensitivity of culture, it is still an arduous process and studies in which *B. burgdorferi* were recovered required the use of large volumes of blood and would therefore be unlikely to be suitable for routine diagnosis (166, 167). A recent study describing a novel method for serum culture claimed 94% sensitivity and 100% specificity (168) lacked adequate steps to rule out false positivity due to contamination. The results of this study have not been able to be validated in other laboratories.

1.6.2. Standard Two-Tier Testing Protocol for Lyme disease diagnosis

The PHE Rare & Imported Pathogens laboratory (RIPL) provides specialist Lyme disease testing for England and Wales, and follows an STT algorithm. The first test uses an ELISA to detect host serum antibodies against the C6 peptide, a conserved 26-amino acid sequence of the variable major protein like sequence-expressed (VLsE) lipoprotein of *B. burgdorferi*. The test has been shown to be capable of detecting all the major infectious European *Borrelia* species (*B. afzelii, B. garinii* and *B.burgdorferi* s.s.) (169). The C6 ELISA measures the presence of both IgG and IgM antibodies and does not differentiate the abundance of one from the other. In the case of a positive or indeterminate result by C6 ELISA, samples are taken forward for confirmatory testing. At PHE RIPL, this takes the form of an array-based immunoblot with separate tests for IgG and IgM antibodies. The base of each well of a 96-well

plate is pre-printed with antigen spots (in triplicate) and uses a number of antigens from *Borrelia,* including native antigens purified from the bacteria together with recombinant VlsE antigens at defined positions, each in triplicate. The interpretation criteria of the tests are dependent on the type (IgG or IgM) and are based upon detection of a requisite number of positive reactions to a defined range of antigens. The two-tier method was developed from 1995 onwards by the CDC and other agencies (170). The fist-step in the assay is intended to generate an objective value that correlates with the degree of the antibody reaction to infection. As the confirmatory second-step is ultimately based on visual interpretation, pairing the step with the objective C6 ELISA serves to reduce erroneous positive results if weak bands are scored as positive (171). **Figure 1.6** shows the pathway used by PHE RIPL for diagnosis of Lyme disease.

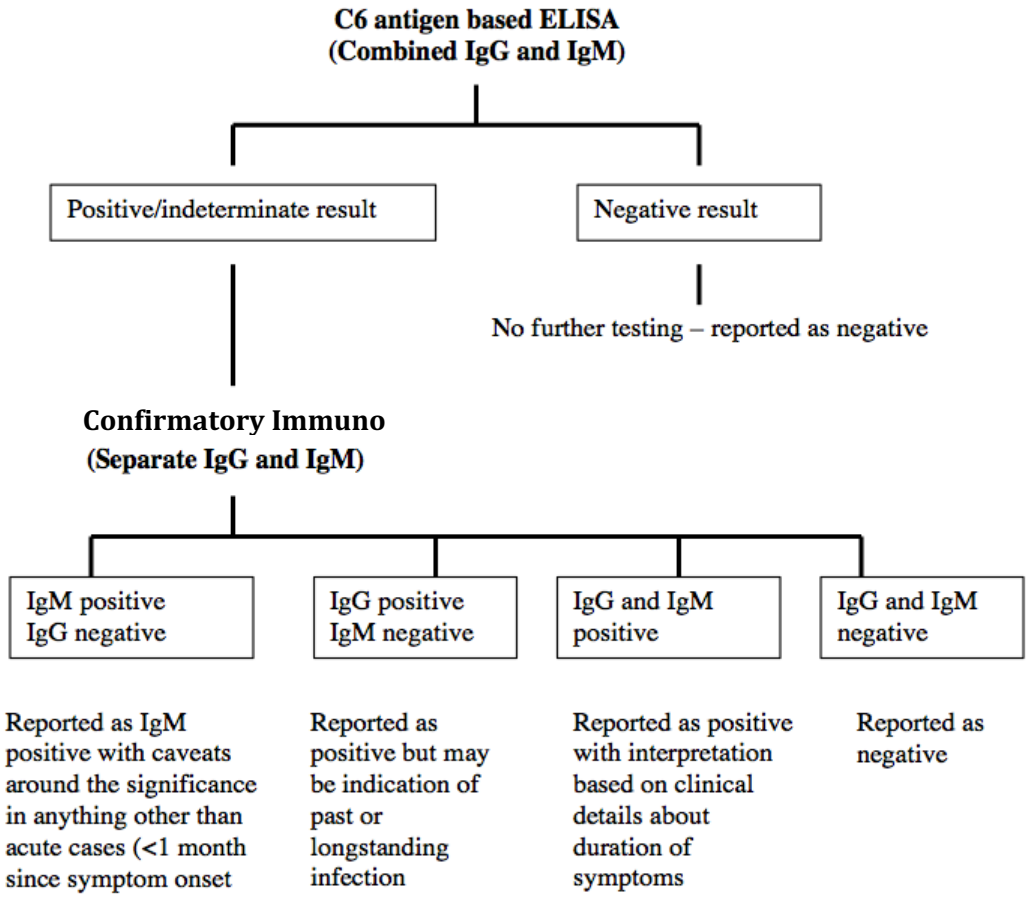


Figure 1.6: Diagram showing the pathway of testing for routine diagnosis of Lyme disease in the England, Wales and Northern Ireland. Serological testing should interpreted in the context of the patient's clinical history

1.6.3. Accuracy of STT testing for Lyme disease. A review of European and North American studies

Previous studies into the efficacy of STT and related serological methods have reported a wide range of values for sensitivity and specificity of said tests. A variety of factors including disease stage, specific assay format and testing regime can influence the result of the test. A broad conclusion based on these studies is that serological testing is less efficacious in early disease when host immune response to infection is still developing. When used in patients with manifestations of disseminated and late disease, including neuroborreliosis and Lyme arthritis,

sensitivity and specificity are much improved. Here, available studies into serological testing for Lyme in both Europe and North America were reviewed and details of the reported sensitivity and specificity of the test were recorded. **Table 1.1-1.4** show a summary of data from available studies into serological testing for Lyme disease, and is divided into subcategories for early Lyme disease, Lyme arthritis, acrodermatitis chronica atrophicans and neuroborreliosis. Reported sensitivity and specificity for IgG and IgM are given where available, together with a quality score based on several factors including cohort size, inclusion of appropriate controls and risk of bias or potential conflict of interest.

1.6.4. Shortcomings of STT testing for Lyme disease diagnosis

Serological testing for Lyme disease requires the detection of antibodies against borrelia antigens. Unsurprisingly, during acute stages of infection, when the adaptive immune response is still developing, serological testing is less effective showing poor sensitivity. A study by Aguero-Rosenfeld (172) used a commercially available ELISA and immunoblot to test sera from 100 patients at various stages of Lyme disease and found that sera of all patients with a history of EM for more than 14 days duration were reactive by both tests. A later study by the same group investigated the appearance of IgM and IgG to borrelial antigens in patients with culture-proven EM. 257 serially collected samples were tested by the 2-tier (STT) method. At baseline, 33% and 43% had positive ELISA and immunoblot results respectively. At 8 to 14 days, 91% of patients had a positive 2-tier result (173). A study by Engstrom *et al* found that the sensitivities of the IgM and IgG immunoblot for detecting seropositive patients were 58.5% and 54.6% respectively at presentation and 100% during their second visit, 8 to 12 days into treatment (174). Patients may therefore test

seronegative for Lyme disease by confirmatory Western blots at time of presentation despite having clinical symptoms of Lyme disease and assessment of in-house scoring criteria for WB analyses is essential (175), together with consideration of clinical symptoms during diagnosis. Due to the persistence of antibodies and the development of memory B and T cells, patients can continue to test positive for Lyme disease long after infection has been cleared by antibiotic treatment.

It has been shown that the genotypes of *B. burgdorferi* s.l. often cause differing manifestations of disease with some having greater proclivities for certain tissues than others (176). A study by Wormser *et al* (177) investigated the effect of genotype on the sensitivity of C6 and STT testing in the USA. Testing was performed on acute-phase serum samples obtained from 158 patients with EM for whom genotype of the isolate was defined based on analysis of ribosomal DNA or genetic variation in OspC. The group found that the sensitivity of C6 testing alone was significantly greater than STT testing. The sensitivity of STT testing varied according to genotype, with one genotype group twice as likely as another to be seropositive despite all patients having culture-confirmed Lyme disease indicating infecting genotype does influence the efficacy of current 2-tier testing protocols. A later study by the same group investigated the utility of diagnostics designed for the US for Lyme disease acquired in Europe and vice versa (178). In the US, Lyme disease is caused exclusively by *Borrelia burgdorferi* s.s. while in Europe, several pathogenic strains circulate including *B. afzelli* and *B. garinii* (171, 179). The study found that the sensitivity of C6 two-tier testing with US assays was 22.5% for detection of LB acquired in Europe. In European patients with neurological involvement, the sensitivity of the US test was

even poorer at 20%. Sensitivity of the European C6 two-tier testing was significantly higher at 70%.

1.6.5. European and North American Systematic reviews of serological testing for Lyme disease

In 2016, the European Centre for Disease Prevention and Control published a systematic literature review on the diagnostic accuracy of serological tests for Lyme disease. The summary estimates for sensitivity for any enzyme immunoassay or immunoblot were: 0.50 (95% CI 0.40-0.61) for erythema migrans stage, 0.77 (95% CI 0.67-0.85) for neuroborreliosis, 0.96 (95% CI 0.89-0.98) for Lyme arthritis, 0.97 (95% CI 0.94-0.99) for acrodermatitis chronica atrophicans and 0.73 (95% CI 0.54-0.87) for unspecified disease stage. Estimates for specificity were around 95% for all disease stages (180). A large heterogeneity was observed for all disease stages and the study concluded that there was not sufficient evidence to make inferences about the value of tests for clinical practice; however it was noted that in early stages of disease, sensitivity and specificity of serological tests were demonstrably lower.

A similar systematic review, also published in 2016, was undertaken by the Public Health Agency of Canada and covered North American research into the accuracy of diagnostic tests for Lyme disease. The summary estimates for sensitivity for any enzyme immunoassay or immunoblot were: 0.46 (95% CI 0.39-0.54) for early acute Lyme disease, 0.90 (95% CI 0.78-0.95) for early neurological and cardiac Lyme disease and 0.99 (95% CI 0.96-0.99) for late neuroborreliosis and Lyme arthritis (181). The wide range of serological testing protocols in laboratories across North America was noted, with validation data available for only a small proportion of licensed assays. A large number of "in house" assays are also in use. For the majority

of these, composition of the test is not publicly available and they are not validated by peer-reviewed literature. Criteria used for the evaluation of serological tests, particularly those that require discretion of the person conducting the test such as Western blot, was also rarely described; therefore, comparison between testing regimes and commercial tests was not particularly informative. Generally, the results of the North American review correlated with those of the European review: a positive association between duration of infection/stage of disease and sensitivity of serological Lyme disease tests. A large heterogeneity between studies was found that could only partially be explained by covariates. Some studies may have used selection criteria that only included patients that were culture positive for Lyme disease. Both systemic reviews concluded that serological testing for Lyme disease needs to be conducted with caution, and should be conducted in combination with considerations of clinical presentation and patient history.

Table 1.1: Sensitivity and specificity of serology based tests for patients with early/EM rash stage Lyme disease

Author/ Date	Quality score	Region	Disease diagnosis	Testing protocol	Serum sample group sizes			Sensitivity (%)			Specificity (%)		
					Lyme patients	Healthy controls	Related-disease controls	IgG	IgM	IgG + IgM	IgG	IgM	IgG + IgM
Hansen *et al.* 1989	2	Europe	Early localised (unclear clinical diagnosis w/ EM)	ELISA (in house)	107	200	98	11.2	16.6	16.6	97.0	95.5	97.0
				ELISA (fla)	107	200	98	35.5	44.8	44.8	95.0	95.8	95.8
Karlsson *et al.* 1989	2	Europe	EM acute stage	ELISA (IgM capture)	30	0	73	-	33	33	-	97.3	97.3
				ELISA (indirect IgM)	30	0	73	-	27	27	-	90.4	90.4
Bergstrom *et al.* 1991	2	Europe	EM acute stage (clinical diagnosis)	In-house EIA	30	64	151	-	-	43	-	-	98.6
Wilske *et al.* 1993	2	Europe	EM (unclear)	IFA (ABS)	31	100	42	45	32	45	100	99	100
				ELISA (OGP)	31	100	42	35	45	45	93	99	99
				ELISA (fla)	31	100	42	39	-	39	93	91.5	93
Rijpkema *et al.* 1994	2	Europe	Early LB (EM, atypical symptoms + tick bite)	Inhibition ELISA (fla)	61	10	31	52.5	6.6	52.5	100	100	100
Engstrom *et al.* 1995	2	N. America	Early Lyme disease (w/ physician-documented EM)	ELISA (whole-cell)	55	75	84	23.6	34.5	45.5	88.7	93.7	97.5
Rauer *et al.* 1995	1	Europe	EM acute stage (clinical diagnosis)	ELISA (p83)	118	154	134	20.3	6	20.3	100	96	100
Hofmann *et al.* 1996	2	Europe	EM acute stage	ELISA (Dako, Behring)	64	57	18	25	40	40	91	93	93
Johnson *et al.* 1996	2	N. America	EM acute stage (clinical diagnosis, culture confirmed)	ELISA (fla) alone	43	113	111	-	-	44	-	-	90.5
				Two-tier protocol (w/ immunoblot)	43	113	111	-	-	56	-	-	95
				ELISA (fla) alone	15	113	111	-	-	53	-	-	90.5
				Two-tier protocol (w/ immunoblot)	15	113	111	-	-	82	-	-	95

Fawcett *et al.* 1998	2	N. America	EM acute stage (clinical diagnosis, early isolated)	Western blot (MarBlot)	10	0	81	40	70	70	-	-	76
				Immundot (GenBio)	10	0	81	10	40	40	-	-	95
Mathiesen *et al.* 1998	3	Europe	EM acute stage (clinical diagnosis)	ELISA (pepC10)	80	100	138	36.3	5	36.3	-	93.6	93.6
				ELISA (fla)	80	100	138	43.8	5	43.8	-	87.7	87.7
				ELISA (rOspC)	80	100	138	37.5	23.8	37.5	-	88	88
Ryffel *et al.* 1998	1	Europe	EM acute stage (clinical diagnosis)	In-house immunoblot	35	180	50	66	49	66	89.6	95.7	89.6
				ELISA (Serion)	22	0	40	83	91	91	92	52	92
				ELISA (C6)	22	0	40	-	-	85	-	-	92
				ELISA (Virotech)	158	200	55	-	-	35	-	-	69.8
				ELISA (Liason)	158	200	55	-	-	37	-	-	89.2
Liang *et al.* 1999	2	N. America	Acute stage, localised (clinical diganosis)	ELISA (C6)	39	0	176	-	-	74	-	-	99
Wilske *et al.* 1999	1	Europe	In-house immunoblot (IgG)	EM (clinical diagnosis)	66	118	21	10.6	-	10.6	97.8	-	97.8
Gomes-Solecki *et al.* 2001	2	N. America	EM acute stage (clinical diagnosis)	ELISA (Non-OspA)	67	10	49	-	-	46	-	-	88.1
				ELISA (WCS)	67	10	49	-	-	43	-	-	98.3
Callister *et al.* 2002	2	N. America	Single EM (clinical diagnosis, IFA positive)	Western blot (MRL diagnostics)	22	0	34	-	-	55	-	-	91.2
Christova *et al.* 2003	2	Europe	EM acute stage (clinical diagnosis w/ typical lesion)	ELISA (Boehring)	105	90	0	17	49	49	100	93	100
Hernandez *et al.* 2003	2	Europe	EM acute stage (clinical diagnosis, localised infection)	Immunoblot (BAG)	42	0	129	20.8	37.5	98.4	98.4	-	98.4
Lahdenne *et al.* 2003	2	Europe	EM acute stage	ELISA (BBK32)	75	40	0	-	87	87	-	99.2	99.2
				ELISA (fla)	75	40	0	-	39	39	-	-	
				ELISA (anti-IR6)	75	40	0	-	39	39	-	-	-
Lovic *et al.* 2003	3	Europe	EM acute stage	IFA	40	80	40	-	-	32.5	-	-	89
				In-house ELISA	40	80	40	-	-	65	-	-	93
				Immunoblot (B. afzelii)	40	80	40	-	-	92.5	-	-	96

Table 1.1 continued.

Bacon 2004	2	N. America	EM acute stage (culture confirmed, clinical diagnosis)	ELISA (rVlsE1)	80	257	302	44	19	44	99	99	99
				ELIA (C6 IgG)	80	257	302	45	-	45	99	-	99
				ELISA (pepC10 IgM)	80	257	302	-	40	40	-	99	99
				Standard two-tiered protocol	80	257	302	-	-	38	-	-	99
Ekerfelt *et al.* 2004	4	Europe	EM acute stage (clinical diagnosis)	ELISA (fla)	17	91	26	23.5	47.5	47.5	96.1	96.1	96.1
				ELISA (Abbot)	17	91	26	11.8	5.9	11.8	93.1	100	100
				ELISA (Mikrogen)	17	91	26	41.2	58.8	58.8	89.2	89.2	89.2
				ELISA (Progen)	17	91	26	58.8	29.4	58.8	78.4	88.2	88.2
				ELISA (Behring)	17	91	26	29.4	47.1	47.1	92.2	86.3	92.2
Goettner *et al.* 2004	1	Europe	EM acute stage (clinical diagnosis, well-defined lesion)	In-house Line immunoblot	15	60	50	53.3	86.7	86.7	99.1	98.2	99.1
Marangoni *et al.* 2005	3	Europe	EM acute stage (culture confirmed)	ELISA (Mikrogen)	45	234	40	57.9	55.7	57.9	97.1	98.9	98.9
				ELISA (Behring)	45	234	40	36.8	70.5	70.5	90.1	92.3	92.3
				ELISA (C6)	45	234	40	-	-	62.1	-	-	96.7
Cinco *et al.* 2006	2	Europe	EM acute stage (culture confirmed)	ELISA (C6)	54	24	0	-	-	63	-	-	100
				ELISA (Enzygnost)	54	0	0	-	-	50	-	-	-
Cerar *et al.* 2006	3	Europe	EM acute stage (Clinical diagnosis, solitary lesion)	EIA (IFA)	76	49	9.2	9.2	1.3	9.2	81.6	98	98
				LIASON	76	49	0	15.8	55.3	55.3	8.6	98	98
Smismans *et al.* 2006	2	Europe	EM acute stage (clinical diagnosis)	ELISA (Dako)	22	0	40	42	61	61	100	78	100
Tjernberg *et al.* 2006	2	Europe	EM acute stage (clinical diagnosis)	ELISA (C6)	158	200	55	-	-	34	-	-	71.2
Jobe *et al.* 2008	2	N. America	EM lesion (clinical diagnosis, significant tick exposure)	ELISA (OspC7)	86			-	-	80	-	-	98
				Western blot (Marblot)	86			-	-	57	-	-	98

Table 1.1 continued.

Ledue *et al.* 2008	2	N. America	Early localised disease (EM, clinical diagnosis, <2 mo duration)	VlsE assay	45	600	207	-	-	64.4	-	-	98
				ELISA (C6)	45	600	207	-	-	68.9	-	-	97.9
				Western blot (MarDx)	45	600	207	-	-	57.8	-	-	-
Lencakova *et al.* 2008	3	Europe	EM acute stage	ELISA (Behring)	54	40	20	42.6	63	63	98.3	98.3	98.3
				IFA	54	40	20	42.1	35.1	42.1	98.3	98.3	98.3
				Immunoblot	54	40	20	53.7	61.1	61.1	98.3	100	100
Marangoni *et al.* 2008	3	Europe	EM acute stage (culture confirmed)	ELISA (Behring)	66	300	100	56.1	54.5	56.1	98.3	95.3	98.3
				ELISA (LIAISON)	66	300	100	39.4	24.3	49.4	96.5	92.8	96.5
Wormser *et al.* 2008	2	N. America	EM acute stage (clinical diagnosis, culture confirmed)	ELISA (C6) alone	158	0	0	-	-	68.4	-	-	-
Branda *et al.* 2010		N. America	EM acute stage (clinical diagnosis, culture confirmed)	Standard two-tiered protocol	33	166	103	-	-	31	-	-	100
				Hybrid IgG two-tiered protocol	33	166	103	-	-	34	-	-	100
Branda *et al.* 2011	3	N. America	EM active stage (clinical diagnosis, culture confirmed)	C6 EIA alone	114	2492	54	-	-	56	-	-	98.4
				Standard two-tiered protocol	114	2492	54	-	-	42	-	-	99.5
				2-EIA algorithm	114	2492	54	-	-	53	-	-	99.5
				Immunoblot (In-house)	55	75	84	43.6	43.6	54.5	88.7	93.7	
Branda *et al.* 2013	2	N. America*	EM acute stage (typical EM, culture-confirmed) *US patients with European-acquired LD	ELISA (C6)	20	100	0	-	-	59.3	-	-	98

Table 1.2: Sensitivity and specificity of serology based tests for patients with Lyme arthritis

Author/ Date	Quality score	Region	Disease diagnosis	Testing protocol	Serum sample group sizes			Sensitivity (%)			Specificity (%)		
					Lyme patients	Healthy controls	Related-disease controls	IgG	IgM	IgG + IgM	IgG	IgM	IgG + IgM
Blaauw *et al.* 1993	3	Europe	LA (clinical diagnosis, severity ranking)	ELISA (In-house)	15	35	0	100	-	100	95	-	95
Wilske *et al.* 1993	2	Europe	LA (clinical diagnosis)	IFA (ABS)	24	100	42	100	4	100	100	99	100
				ELISA (OGP)	24	100	42	100	29	100	93	99	99
				ELISA (fla)	24	100	42	92	17	92	93	91.5	93
Johnson *et al.* 1996	2	N. America	LA (clinical diagnosis, intermittent attacks of oligoarticular swelling)	ELISA (fla) alone	36	113	111	-	-	89	-	-	90.5
				Two-tier protocol (w/ immunoblot)	36	113	111	-	-	100	-	-	95
Liang *et al.* 1999	2	N. America	LA (clinical diagnosis)	ELISA (C6)	49	0	176	-	-	100	-	-	99
Bacon *et al.* 2003	3	N. America	LA (intermittent objective swelling of one or more large joints	ELISA (rVlsE1)	33	257	302	97	39	97	99	99	99
				ELIA (C6 IgG)	33	257	302	94	-	94	99	-	99
				ELISA (pepC10 IgM)	33	257	302	-	39	39	-	99	99
				Two-tiered protocol	33	257	302	-	-	97	-	-	99
Goettner *et al.* 2005	1	Europe	LA (unclear/ clinical diagnosis)	In-house line immunoblot	10	60	50	100	60	100	99.1	98.2	99.1
Cinco *et al.* 2006	2	Europe	LA (clinical diagnosis)	ELISA (C6)	15	24	0	-	-	100	-	-	100
				ELISA (Enzygnost)	15	0	0	-	-	100	-	-	-

Lencakova *et al.* 2008	3	Europe	LA (clinical diagnosis, oglioarthritis of large joints)	ELISA (Behring)	13	40	20	92.3	0	92.3	98.3	98.3	98.3
				IFA	13	40	20	76.9	0	76.9	98.3	98.3	98.3
				Immunoblot	13	40	20	100	0	100	98.3	100	100
Branda *et al.* 2013	2	Europe		ELISA only (European)	15	100	0	100	60	100	98	98	98
				ELISA only (N. America)	15	100	0	-	-	100	-	-	97
				Conventional 2-tier (European)	15	100	0	-	-	93	-	-	99
				Conventional 2-tier (N. America)	15	100	0	-	-	60	-	-	100

Table 1.3: Sensitivity and specificity of serology based tests for patients with Lyme Acrodermatitis Chronica Atrophicans

Author/ Date	Quality score	Region	Disease diagnosis	Testing protocol	Serum sample group sizes			Sensitivity (%)			Specificity (%)		
					Lyme patients	Healthy controls	Related-disease controls	IgG	IgM	IgG + IgM	IgG	IgM	IgG + IgM
Hansen *et al.* 1989	2	Europe	ACA (clinical diagnosis, histopathology)	ELISA (in-house)	50	200	98	98	22	98	97	95.5	97
				ELISA (fla)	50	200	98	100	12	100	95	95.8	95.8
Karlsson *et al.* 1989	3	Europe	ACA (clinical diagnosis, histopathology)	ELISA (IgM capture)	10	0	73	-	0	0	-	97.3	97.3
				ELISA (indirect IgM)	10	0	73	-	30	30	-	90.4	90.4
Bergstrom *et al.* 1991	2	Europe	ACA (clinical diagnosis)	In-house EIA	22	64	151	-	-	100	-	-	98.6
Wilske *et al.* 1993	2	Europe	ACA (clinical diagnosis)	IFA (ABS)	19	100	42	100	5	100	100	99	100
				ELISA (OGP)	19	100	42	100	16	100	93	99	99
				ELISA (fla)	19	100	42	84	11	84	93	91.5	93
Rauer *et al.* 1995	1	Europe		ELISA (p83)	17	154	134	94.1	11.8	94.1	100	96	100
Hofmann *et al.* 1996	2	Europe	ACA (unclear/ clinical diagnosis)	ELISA (Dako)	31	57	18	100	26	100	85.3	90.3	90.3
				ELISA (Behring)	31	57	18	100	26	100	92.5	92.5	92.5
Ryffel *et al.* 1997	1	Europe	ACA (clinical diagnosis)	In-house immunoblot	27	180	50	96	9	96	89.6	95.7	95.7
Goettner *et al.* 2004	1	Europe	ACA (unclear/ clinical diagnosis)	In-house line immunoblot	10	60	50	100	70	100	99.1	98.2	99.1
Branda *et al.* 2013	2	Europe	ACA (clinical diagnosis w/ supportive histologic findings	ELISA only (European)	14	100	0	100	71	100	98	98	98
				ELISA only (N. America)	14	100	0	-	-	100	-	-	97
				Conventional 2-tier (European)	14	100	0	-	-	100	-	-	99
				Conventional 2-tier (N. America)	14	100	0	-	-	100	-	-	100

Table 1.4: Sensitivity and specificity of serological tests for Lyme disease: Neuroborreliosis

Author/ Date	Quality score	Region	Disease diagnosis	Testing protocol	Serum sample group sizes			Sensitivity (%)			Specificity (%)		
					Lyme patients	Healthy controls	Related-disease controls	IgG	IgM	IgG + IgM	IgG	IgM	IgG + IgM
Karlsson *et al.* 1989	2	Europe	NB (CSF pleocytosis, clinical diagnosis)	ELISA (IgM capture)	37	0	73	-	37.8	37.8	-	97.3	97.3
				ELISA (indirect IgM)	37	0	73	-	54	54	-	90.4	90.4
Wilske *et al.* 1993	2	Europe	NB (CSF lymphocytic pleocytosis, positive antibody index	IFA (ABS)	24	100	42	100	4	100	100	99	100
				ELISA (OGP)	24	100	42	100	29	100	93	99	99
				ELISA (fla)	24	100	42	92	17	92	93	91.5	93
Rauer *et al.* 1995	1	Europe	NB (clinical diagnosis - stage II and II, late - chronic)	ELISA (p83)	118	154	134	100	21.4	100	100	96	100
Johnson 1996	2	N. America	Early and late neurologic (clinical diagnosis, meningitis, polyneuropathy etc.)	ELISA (fla) alone	17	113	111	-	-	94.1	-	-	90.5
Mathiesen *et al.* 1998	3	Europe	NB (CSF pleocytosis, intrathecal antibody index)	ELISA (pepC10)	100	100	138	8	45	45	-	93.6	93.6
				ELISA (fla)	100	100	138	52	63	63	-	87.7	87.7
				ELISA (rOspC)	100	100	138	6	48	48	-	88	88
Ryffel *et al.* 1998	1	Europe	NB (clinical diagnosis)	In-house immunoblot	61	180	50	75	52	75	89.6	95.7	95.7
Liang *et al.* 1999	2	N. America	Early neuroborreliosis (clinical diagnosis)	ELISA (C6)	20	0	176	-	-	95	-	-	99
			Late neuroborreliosis	ELISA (C6)	10	0	176	-	-	100	-	-	100

Bacon *et al.* 2003	2	N. America	Late NB (at least 1 manifestation of late NB and laboratory confirmation of infection)	ELISA (rVlsE1)	11	257	302	100	9	100	99	99	99
				ELIA (C6 IgG)	11	257	302	73	-	73	99	-	99
				ELISA (pepC10 IgM)	11	257	302	-	18	18	-	99	99
				Two-tiered protocol	11	257	302	-	-	100	-	-	99
Goettner *et al.* 2005	1	Europe	NB (clinical diagnosis - stage II, intrathecal antibody production index)	In-house Line immunoblot	50	60	50	-	-	88	-	-	99.1
				Western blot (In-house)	50	60	50	-	-	72	-	-	99.1
Cerar *et al.* 2006		Europe	NB (clinical diagnosis)	EIA (IFA)	28	49	0	85.7	7.1	85.7	81.6	98	98
				LIASON	28	49	0	71.4	3.4	71.4	8.6	98	98
Cinco *et al.* 2006	3	Europe	NB (unclear clinical diagnosis)	ELISA (C6)	6	24	0	-	-	100	-	-	100
				ELISA (Enzygnost)	6	0	0	-	-	100	-	-	100
Tjernberg *et al.* 2006	2	Europe	NB (clinical diagnosis, positive CFS antibody index)	ELISA (C6, Virotech Liason)	26	200	55	-	-	85.3	-	-	76.6
Lencakova *et al.* 2008	3	Europe	NB (clinical diagnosis, stage II, stage III, CSF lymphocytic pleocytosis)	ELISA (Behring)	7	40	20	57.1	42.9	57.1	98.3	98.3	98.3
				IFA	7	40	20	14.3	14.3	14.3	98.3	98.3	98.3
				Immunoblot	7	40	20	57.1	28.6	57.1	98.3	100	100

Table 1.4 continued.

Branda *et al.* 2011	3	N. America	Late disease (late neuritis or arthritis, clinical diagnosis)	C6 EIA alone	29	2492	54	-	-	100	-	-	98.4
				Standard two-tiered protocol	29	2492	54	-	-	100	-	-	99.5
				2-EIA algorithm	29	2492	54	-	-	100	-	-	99.5
				Two-tier protocol (w/ immunoblot)	17	113	111	-	-	100	-	-	95
Branda *et al.* 2013	3	Europe	NB (CSF pleocytosis w/ EM, isolation of Borrelia or positive antibody index)	ELISA only (European)	15	100	0	87	80	87	98	98	98
				ELISA only (N. America)	15	100	0	-	-	87	-	-	97
				Conventional 2-tier (European)	15	100	0	-	-	87	-	-	99
				Conventional 2-tier (N. America)	15	100	0	-	-	40	-	-	100

1.6.6. Novel testing methods for diagnosis of Lyme disease

Although serology is the mainstay of Lyme disease laboratory diagnosis by state run laboratories across the world, a wide range of alternative tests are offered by private laboratories particularly in the US and Europe. The following discussion focuses on a critique of T-cell based tests and tests for antigen detection in urine.

1.6.6.1. T-cell directed tests

The Lymphocyte Transformation Test (LTT) measures the proliferation of antigen specific T cells in response to antigen challenge *in vitro*, from which it is assumed that proliferation correlates with previous sensitisation of T-cells to the antigen. Peripheral blood mononuclear cells are cultured with different concentrations of the antigen of interest for a period of several days. Throughout the 1990s, several groups investigated the potential of LTTs for the diagnosis of Lyme disease (182-185) mostly using lysates from whole *B. burgdorferi* that were antigenically heterogeneous and included known non-specific mitogens such as bacterial lipopolysaccharides (186, 187). The LTTs themselves were often non-validated and non-standardised and the sensitivity and specificity were found to be generally poor in most studies and results were not reproducible.

In 2006, Valentine-Thon *et al* reported an improved LTT format, the memory lymphocyte immunostimulation assay (LTT-MELISA) for use in the diagnosis of Lyme disease (186). The test uses a higher number of lymphocytes ($1x10^6$ per well) with 4 well-characterised recombinant *Borrelia* antigens (from *B. afzelii* and *B. garinii* but not *B. burgdorferi* s.s.) as stimuli. Similarly, Von Baehr *et al* also sought

to improve the borrelia LTT by using lysate antigens from *B. burgdorferi* s.s., *B. afzelii* and *B. garinii* and recombinant OspC (188).

Although LTTs may show some utility for monitoring Lyme disease progression and resolution following antibiotic treatment (123) there is insufficient evidence to support their use as a primary diagnostic tool for the differentiation of active disease (124). Independent studies are needed using well-characterised cases and controls to determine whether LTTs have any clinical validity for the diagnosis of Lyme disease. An alternative type of test that detects stimulated antigen specific T cells is the enzyme-linked immunospot (ELISpot). After short, typically overnight, antigen challenge, cytokine secretion, commonly interferon-gamma (IFN-γ), by activated antigen-specific T cells is detected as a spot around the secreting cell. These tests can also be known as IFN-γ release assays (IGRAs) and IFN-γ can be measured in the soluble phase as an alternative to the spot-based ELISpot systems. Commercial IGRA and ELISPOT test are licensed for the diagnosis of tuberculosis.

As with LTT, there is great variability in ELISpot formats, particularly around the selection of stimulating antigens and questions surround the clinical validity of ELISpot- based tests for Lyme disease. There is again an urgent need for independent studies to determine whether ELISpots have a higher sensitivity during early infection than STTT or MTTT and whether they are able to discriminate between past and active infection. Preliminary evidence suggests that the ELISpot cannot discriminate between past and active neuroborreliosis (189) and more comprehensive studies are underway to address these questions (190).

1.6.6.2. Alternative tests for Lyme disease

Another proposed method for Lyme disease diagnosis is the detection of biological material from *B. burgdorferi* in the host. Many of the tests based on this method are likely to suffer the same issues seen in the use of PCR for detection of bacterial DNA: that in Lyme disease, bacteraemia is low and any testing methods would need to be very sensitive, assuming any bacterial material is present in a test sample at all. An antigen capture test for OspA of *B. burgdorferi* in patient urine has been proposed several times since the mid-1990s, but has repeatedly been shown to be unreliable in peer-reviewed validation studies (191). A 2015 study into the use of particle based high-affinity chemical baits, known as nanotrap, for detection of OspA in urine showed sensitivity and specificity values approaching 100% for early Lyme disease (192); however, the study was funded by the test manufacturer and no independent studies assessing the approach have subsequently been published. Several other proposed testing methods for Lyme disease including measurement of antibodies in joint fluid, quantitative CD57 lymphocyte assays and immunofluorescence staining lack adequate scientific review and are not recommended for diagnosis of Lyme disease by the CDC (193).

1.7. PROTEOMIC ANALYSIS FOR DIAGNOSTIC BIOMARKER DISCOVERY

Proteomics is the large-scale study of a set of proteins produced in an organism, system or within a given biological specimen. Advances in study methods and instruments has allowed the field to move from being purely qualitative to being quantitative, allowing comparison of the quantity of specific proteins between sample groups with increasing accuracy. This has led to great interest into the discovery of proteins that are biomarkers of disease; proteins that are consistently present at abnormal levels that can be measured with specific and sensitive methods. Over the last two decades, many studies have utilised proteomics approaches for discovery of potential biomarkers for many diseases, including genetic disorders, cancer and infection.

1.7.1. Proteomic based studies of Lyme disease

Relatively few proteomic based Lyme disease biomarker discovery studies have been published. A 2011 study by Schutzer *et al* (194) used a high-resolution mass spectrometry approach to compare pooled cerebrospinal fluid from well-characterised Neurologic Post Treatment Lyme disease (nPTLS) patients, Chronic Fatigue Syndrome (CFS) patients and healthy controls. They found that groups and individuals within groups could be distinguished based on their CSF proteins. While these comparisons were made between the entire CSF proteomes of the groups, the results showed that each condition has a multitude of candidate diagnostic biomarkers for future validation studies. A similar study was conducted in 2012, comparing the CSF proteome of patients with acute Lyme disease to a control group (171). Liquid

chromatography tandem mass spectrometry (LS-MS/MS) was used to compare pooled CSF patients diagnosed with early disseminated Lyme disease and CSF inflammation (n=26) against control samples (n=19) and identified 108 proteins that differed significantly and consistently between groups.

Both of these studies showed promising results and proof of concept that highly sensitive proteomic analysis can be used to distinguish between patient groups. However, it should be noted that both studies were based on the analysis of cerebrospinal fluid that, while representing the best biological specimen for studying the pathogenesis of brain infections, including Lyme neuroborreliosis (LNB), would not form the ideal basis for a novel diagnostic test for early Lyme disease The lumbar puncture operation involved in obtaining CSF is not a routine procedure and is usually avoided unless neurological complications are suspected (195). In a targeted approach of 58 immune mediators and 7 acute phase proteins, the T cell chemokines CXCL9, CXCL10 and CCL19 and serum amyloid A1 were found to be increased the serum of Lyme disease patients (196). A similar targeted study of chemokine response proposed the protein CCL19 as a marker of posttreatment Lyme disease syndrome (197).

1.7.2. Proteomic biomarker discovery in other infectious diseases

While there has been relatively little research into proteomics-based biomarker discovery in LD, many studies have been conducted on several other infectious diseases. The decision to treat active tuberculosis (TB) shares several problems with Lyme disease diagnosis. In a similar manner to the 2-tier serology based test currently in use for LD, the tuberculin skin test and interferon-gamma release assay are based

on the host immune response and cannot distinguish active TB from cleared or latent infection (198, 199). As with *Borrelia*, mycobacteria are slow growing bacteria and attempts to culture from patient samples have low sensitivity (200). DNA detection by PCR is often complicated due to difficulty in obtaining suitable samples from the site of infection. Microscopy remains the first-line diagnostic approach, but its efficacy is dependent on bacteremia and ability of the operator. Therefore, many studies have investigated the discovery of biomarkers for improved diagnostics of active TB. A 2013 study by Liu *et al* (201) comparing serum samples from 180 cases of TB to 211 control samples established a model based on four biomarker proteins that could distinguish TB from controls with sensitivity and specificity of 83.3% and 84.2% respectively.

While studies such as this and numerous others have shown promising results, there has been relatively little progress in the development of biomarker based diagnostic tests. A review by Haas *et al* (202) of all TB proteomic biomarker studies published to date found that biomarker candidates reported varied considerably between studies and that there was an inability to verify certain biomarkers in subsequent studies.

1.7.3. Proteomic biomarker discovery in non-infectious diseases

The field of cancer diagnostics has demonstrated the potential of biomarkers. Over-expression of HER-2/neu has previously been shown to play a role in the development of node-negative breast cancer (203, 204). The protein is now used as the basis of several diagnostic tests to determine HER-2 status, with a positive result leading to HER-2 targeting therapy such as trastuzumab. Similarly, overexpression of epidermal growth factor receptor (EGFR) has previously been used as a selection

criterion for the enrolment of patients for clinical studies of the colorectal cancer drug cetuximab (205). In general, cancer biomarkers represent proteins that are involved in disease development and are typically expressed at such high levels to allow detection. This, in part, explains the relative success of biomarkers in cancer diagnostic compared to other disease groups, where biomarkers may represent proteins that are indirectly related to pathogenesis and are therefore more likely to be subject to variability between patients.

1.7.4. Other considerations in proteomic biomarker discovery

The problems faced in the above examples are common to many proteomic based biomarker discovery studies. The lack of reproducibility between studies is likely to be due to several factors including differences in experimental design, patient cohort and inherent variability due to machine limitations and the complexity of biological samples(206). Serum, cerebrospinal fluid and other biological fluid proteomes are complex, with wide protein concentration range and dynamics. In blood, more than 10 orders of magnitude in concentration separate albumin from the rarest proteins present (207). To overcome this complexity, fractionation methods including column chromatography and the use of depletion columns specifically designed to remove highly abundant proteins can be used. While this allows the study of rarer proteins that are more likely to be influenced by infection, the complexity of the process and use of different fractionation and depletion methods contributes to the lack of reproducibility between studies (208).

With the discovery-based 'shotgun' proteomics studies previously described, as many proteins as possible were identified in order to discover potential new biomarkers.

Another approach is the use of targeted proteomics where a select few proteins of interest are analysed in samples with high sensitivity, reproducibility and reliable quantification. A commonly used method in target proteomics is selected reaction monitoring (SRM). SRM commonly utilises triple quadrupole mass spectrometers, with the first stage involving isolation of the precursor ion of interest before further fragmentation to yield product ions which form the mass spectra (209). Using isotopic labeling as a concentration standard, SRM can be also used to provide highly accurate quantification (209). As a more specific approach, targeted proteomics has been shown to be more reproducible between studies than shotgun methods.

1.8. TRANSCRIPTOMIC AND METABOLOMIC ANALYSES FOR BIOMARKER DISCOVERY

1.8.1. Transcriptomic based study of disease

Transcriptomics is the study of the transcriptome – the set of all messenger RNA present in an organism, single cell or biological sample. Many studies examining the human transcriptional response to various bacterial infections have been published. The majority of these were designed to investigate the pathogenesis and immunological response to specific infections rather than being specific biomarker discovery studies; however, with the advent of more sensitive techniques and application of algorithms to identify characteristics including machine-learning, many researchers are now looking to transcriptomic data to identify transcriptomic patterns that may have diagnostic or prognostic potential.

1.8.1.1. Transcriptomic based study of Lyme disease

The first, and currently only, transcriptome analysis of the human host response to acute Lyme disease was published in 2015 (210). The study involved longitudinal transcriptome analysis of PBMCs from 29 Lyme disease patients and 13 matched controls. Transcriptome profiling by RNA sequencing and pathway analysis of samples at three time points ranging from time of Lyme disease diagnosis to 6 months after completion of antibiotic therapy, revealed distinct differences between study groups. As the study involved analysis of the same patients at different stages of disease, comparisons could be made between subsets of patients with differing clinical outcomes including full recovery at 6 months or development of persistent symptoms post-treatment. The underlying pathogenesis and patient predisposition to development of persistent symptoms after treatment, also known as post-treatment Lyme disease syndrome (PTLDS), is currently poorly understood. Interestingly, the study found that the average duration of acute illness i.e. the time of onset of symptoms such as EM and initiation of antibiotic treatment was significantly longer in patients who went on to develop persistent symptoms; suggesting delayed antibiotic treatment may have a role in the development of persistent symptoms and highlights the need for Lyme disease diagnostics capable of detecting the earliest stages of disease. Analysis of Lyme disease patients at time of diagnosis and normal healthy controls revealed a total of 1,235 differentially expressed genes (DEGs). 62% of DEGs identified at the first time point were also differentially expressed at the time of completion of antibiotic therapy. Notably, no differential gene expression was observed between Lyme disease patients with resolved illness after treatment and those that developed persistent symptoms at 6 months. The Lyme disease transcriptome did not fully return to baseline relative to controls even at the final time

point of 6 months showing that the host-response to Lyme disease appears to be active even after clinical resolution of symptoms. Pathway analyses revealed the activation of genes involved with the inflammatory response. Predicted canonical pathways included the elongation initiation factor 2 (eIF2) signaling pathway involved in translation initiation in response to stress during infection. Immune cell trafficking and hematologic system pathways were also found to be increased in Lyme disease positive patients; future studies may therefore benefit from the inclusion of related disease control groups to elucidate how specific the identified DEGs were to early Lyme disease rather than a generic response to acute bacterial infection. The group did go on to compare the RNA-seq data to 12 previously published transcriptome data sets for several other infections other than Lyme disease and found that while a proportion of DEGs were common to several groups, the Lyme disease transcriptome was still distinguishable; however, many of the transcriptomes included used different methodologies or were based on *in vitro* infection of cells and may therefore not be ideal for direct comparison. In a study of the transcriptional patterns seen in EM biopsy samples from untreated adult Lyme disease patients, 254 differentially regulated genes were identified and were characterized by the induction of chemokines, cytokines, antimicrobial peptides and other innate and adaptive immunity related genes (211).

1.8.1.2. Transcriptomic studies in other diseases

While very few transcriptomics papers exist for LD, for other infectious diseases that have been more widely researched, several problems exist despite some promising results. A review of transcriptomic tuberculosis studies by Haas *et al* found that published signatures varied widely in size and showed limited overlap between

studies (212). It is likely that differences in study design, as well as patient selection form the major contributing factor for the lack of continuity between studies and therefore are an important consideration in the design of future transcriptomic studies. Gene transcripts do not necessarily correlate to measurable changes in proteins due to transcripts being non-protein coding, short protein half-life, degradation of mRNA and post-translation modification of proteins(213). A 2012 review found the correlation to be as little as 40% in some systems(214). This is an important consideration when transcription data is to be used to identify potential biomarkers for follow-up work at the protein level.

1.8.2. Metabolomic based study of disease

The metabolome represents the total number of metabolites present within an organism, cell or biological sample. While transcriptomic and proteomic analysis provide information on the set of gene products either being expressed or present in a sample, metabolomics involves the study of small molecule metabolites present. As with many transcriptomics studies, much of the research into the metabolome in infectious disease has the main aim of gaining novel biological insights into pathogenesis; however, as the metabolome is known to be highly influenced by acute infection, it is an attractive area for diagnostic biomarker discovery.

1.8.2.1. Metabolomic based study of Lyme disease

A 2014 study by Molins *et al*. used retrospective serum samples to develop a metabolic biosignature for the detection of early active Lyme disease (215). Well-characterised samples from patients with early Lyme disease were analysed using liquid-chromatography-mass spectrometry and the results compared to a large cohort

of non-Lyme disease controls including patients with related diseases including syphilis, periodontitis, infectious mononucleosis and fibromyalgia. The initial liquid-chromatography stage allowed the analysis of low molecular mass (<1500 Da) biological molecules only and therefore decreased the overall complexity of the sample while allowing high-accuracy detection of small molecules. Sera and corresponding LC-MS data were randomly separated into discovery/training and test-samples. A molecular feature extractor algorithm was then used to develop a metabolic biosignature for early Lyme disease A biosignature of 95 MFs (molecular features) that were consistently increased or decreased was developed for early Lyme disease when compared to healthy controls. Predictive chemical formulae were developed for the majority of the MFs were found and identified metabolites could be classified as lipids or polyunsaturated fatty acids. While the exact origin or function of the metabolites identified is difficult to determine, this is not vital information as the nature of a test based on metabolomic data would only require that the metabolic biosignature would be consistent and reproducible between samples and specific to early Lyme disease. A study of metabolites in the urine of Lyme disease patients in comparison to infectious mononucleoisis patients and normal healthy controls revealed dysregulation of several metabolic processes in early Lyme and showed that metabolic biosignatures could be discriminatory between groups (216).

1.8.2.1. Metabolomic based study of other diseases

Several studies have investigated the potential of metabolic profiling for the diagnosis of infectious disease. A notable example is a 2013 study by Tritten *et al* which identified urinary metabolites as candidate diagnostic markers of malaria. The group used a nuclear magnetic resonance approach followed by multivariate modelling to

identify two urinary metabolites that were found only in *Plasmodium* infected mice. The metabolites have not been described in the mammalian or parasite metabolome to date and were present only in *Plasmodium* infected mice. Antigen-detecting rapid diagnostic tests (RDTs) for malaria have been rigorously studied and tested over recent years and show many advantages over existing tests; however, clinical sensitivity remains relatively poor and high potential of co-infection with parasitic worms complicates diagnosis. Potential co-infection and seropositivity due to previous infection are issues that also affect Lyme disease diagnosis and need to be considered during the development of novel diagnostics.

1.9. Aims of the study

The negative-predictive value of current serological testing protocols for Lyme disease including STT have been questioned, particularly in the earliest stages of disease. In this study, proteomic (mass-spectrometry based) and transcriptomic (RNA-seq based) approaches were taken to attempt to identify host-derived markers that may hold diagnostic potential. Particular interest was placed in the discovery of markers that were detectable in early disease. In addition, computer analyses of proteomic and transcriptomic data were used to further explore the host response to *Borrelia* s.l. infection. The aims of this study are:

1. To investigate the issues surround current testing methodology for Lyme disease.
2. To use label-free quantitative mass-spectrometry and downstream analyses to identify protein markers of Lyme disease.
3. To use mRNA sequencing and differential gene expression analyses to identify gene markers of Lyme disease.
4. To compare early and late stage Lyme disease cases by mRNA sequencing and identify gene markers that are indicative of early disease.
5. To further investigate the host response to *Borrelia* s.l. infection by computer analyses of proteomic and transcriptomic data

CHAPTER 2: MATERIALS AND METHODS

2.1. HUMAN CLINICAL SAMPLES

Human serum, plasma and whole-blood samples were used during this study. The samples are divided into 9 groups based on their infectious disease status and provider.

2.1.1. Ethics statement

All samples were anonymised before receipt. Sample **groups 1, 2, and 5** were provided by the Rare and Imported Pathogen Lab - PHE Porton, and the Bacterial Reference Department – PHE Colindale. They consist of residual samples from routine diagnostic testing. All patients had given consent for anonymised residual samples to be used for research purposes by PHE. **Groups 3, 4 and 9** were provided by Ales Chrdle, Hospital Ceske Budejovice and had prior ethical approval from Czech Republic regulatory bodies for use in research. **Group 7** were provided by the NHS Blood Transfusion Service (NHS-BT) with ethical approval for their use in research in place. **Group 8** were also provided by PHE, Porton and consisted of healthy/normal control samples. Ethical approval for the use human samples in this study as a whole was obtained from the NRES Committee East-Midlands **(appendix A).**

2.1.2. Sample list

GROUP 1: UK LYME DISEASE POSITIVE

Sera submitted for Lyme disease testing that are seropositive and have a clinical history / symptoms consistent with early Lyme disease.

Type: Serum, Retrospective; residual diagnostic samples

Source: RIPL, PHE Porton **Table 2.1:** Group 1 sample details

ID	Sex	Age	Presenting symptoms	C6 Lyme Index	C6 inter-pretation	Virastripe IgM	Virastripe IgG
LP-1	Female	10	EM rash, memory of tick bite	1.032	Positive	Positive	Negative
LP-2	Female	46	EM rash	7.336	Positive	Positive	Positive
LP-3	Male	10	EM rash	8.101	Positive	Positive	Indeterminate
LP-4	Female	56	EM rash	8.325	Positive	Positive	Negative
LP-5	Female	44	EM rash, memory of tick bite	3.049	Positive	Positive	Positive
LP-6	Female	63	EM rash	1.731	Positive	Positive	Negative
LP-7	Male	62	EM rash, memory of tick bite	5.952	Positive	Positive	Positive
LP-8	Male	55	EM rash	3.933	Positive	Positive	Negative
LP-9	Male	47	EM rash, memory of tick bite	8.024	Positive	Positive	Indeterminate
LP-10	Male	54	EM rash	2.353	Positive	Positive	Indeterminate
LP-11	Male	55	EM rash	8.380	Positive	Positive	Positive
LP-12	Female	69	EM rash, systemically unwell	6.015	Positive	Positive	Indeterminate
LP-13	Male	27	EM rash	1.433	Positive	Positive	Positive
LP-14	Female	14	EM rash, memory of tick bite	3.656	Positive	Positive	Positive
LP-15	Female	25	EM rash	1.097	Positive	Positive	Negative
LP-16	Female	61	EM rash, memory of tick bite	1.654	Positive	Positive	Indeterminate
LP-17	Male	44	EM rash	3.022	Positive	Positive	Indeterminate
LP-18	Male	38	EM rash, memory of tick bite	5.011	Positive	Positive	Positive
LP-19	Female	23	EM rash	2.763	Positive	Positive	Positive
LP-20	Male	49	EM rash, memory of tick bite	1.984	Positive	Positive	Negative
LP-21	Female	44	EM rash	1.945	Positive	Positive	Indeterminate
LP-22	Female	56	EM rash	3.062	Positive	Positive	Indeterminate
LP-23	Male	30	EM rash	1.776	Positive	Positive	Positive
LP-24	Male	44	EM rash	2.054	Positive	Positive	Positive
LP-25	Female	71	EM rash	5.012	Positive	Positive	Negative
LP-26	Male	44	EM rash, memory of tick bite	3.041	Positive	Positive	Positive

GROUP 2: UK LYME DISEASE SERONEGATIVE

Sera submitted for Lyme disease testing that are seronegative and have a clinical history / symptoms not consistent with early Lyme disease.

Type: Serum, Retrospective; residual diagnostic samples

Source: RIPL, PHE Porton **Table 2.2:** Group 2 sample details

ID	Sex	Age	Presenting symptoms	C6 Lyme Index	C6 inter-pretation	Virastripe IgM	Virastripe IgG
LN-1	F	20	Unknown	0.494	Negative	N/A	N/A
LN-2	F	62	Unknown	0.390	Negative	N/A	N/A
LN-3	M	9	Unknown	0.354	Negative	N/A	N/A
LN-4	F	30	Headache	0.464	Negative	N/A	N/A
LN-5	F	49	Unknown	0.159	Negative	N/A	N/A
LN-6	M	53	Unknown	0.375	Negative	N/A	N/A
LN-7	M	42	Unknown	0.362	Negative	N/A	N/A
LN-8	F	68	Unknown	0.343	Negative	N/A	N/A
LN-9	F	54	Unknown	0.134	Negative	N/A	N/A
LN-10	M	26	Fatigue	0.126	Negative	N/A	N/A
LN-11	F	22	Unknown	0.173	Negative	N/A	N/A
LN-12	F	59	Unknown	0.132	Negative	N/A	N/A
LN-13	U	24	Unknown	0.238	Negative	N/A	N/A
LN-14	M	53	Unknown	0.300	Negative	N/A	N/A
LN-15	F	44	Unknown	0.405	Negative	N/A	N/A
LN-16	M	46	Unknown	0.273	Negative	N/A	N/A
LN-17	M	58	Unknown	0.213	Negative	N/A	N/A

GROUP 3: CZECH LYME DISEASE POSITIVE

Sera submitted for Lyme disease testing that are seropositive and have a clinical history / symptoms consistent with early Lyme disease.

Type: Serum, Retrospective; residual diagnostic samples

Source: Hospital Ceske Budejovice **Table 2.3:** Group 3 sample details

ID	Sex	Age	Additional clinical notes	C6 Lyme Index	C6 interpretation	Virastripe IgM	Virastripe IgG
CLP-1	Male	58	Skin xerois, eczema numulare	Unknown	Positive	Negative	Positive
CLP-2	Male	42	Generally unwell, arthralgia	Unknown	Positive	Positive	Positive
CLP-3	Female	49	EM	Unknown	Positive	Positive	Positive
CLP-4	Female	60	Fatigue (5 years)	Unknown	Positive	Borderline	Positive
CLP-4(ii)	Female	60	Fatigue (5 years)	Unknown	Positive	Positive	Positive
CLP-5	Male	44	EM	Unknown	Positive	Positive	Positive
CLP-6	Female	48	EM	Unknown	Positive	Positive	Positive
CLP-7	Female	61	Arthralgia, fatigue	Unknown	Positive	Positive	Negative
CLP-8	Male	63	Lyme arthritis	Unknown	Positive	Positive	Positive
CLP-9	Male	45	EM	Unknown	Positive	Negative	Positive
CLP-10	Male	58	Muscle pain, EM, lymphocytoma	Unknown	Positive	Positive	Positive
CLP-11	Female	63	Generally unwell, fatigue	Unknown	Positive	Positive	Negative
CLP-12	Male	41	Lyme knee artritis, synovectomy	Unknown	Positive	Negative	Positive
CLP-13	Male	69	Unknown	Unknown	Positive	Positive	Positive
CLP-14	Female	11	Acute psychosis, Lyme in CSF Neg.	Unknown	Positive	Positive	Positive
CLP-15	Female	7	Unknown	Unknown	Positive	Positive	Positive
CLP-16	Female	18	Arm swelling, arthralgia, onset EM, generally unwell	Unknown	Positive	No test	no test
CLP-17	Male	61	Unknown	Unknown	Positive	Positive	Positive
CLP-19	Male	76	Polyarthritis, mycoplasma pneumonia	Unknown	Positive	Negative	Positive
CLP-20	Female	30	EM	Unknown	Positive	Positive	Positive
CLP-21	Male	62	Hyperkeratosis of the lower limbs	Unknown	Positive	Negative	Positive
CLP-22	Male	63	Rapidly progressing Alzheimer	Unknown	Positive	Negative	Positive
CLP-23	Male	64	Severe low back pain	Unknown	Positive	Positive	Positive
CLP-24	Male	37	Acute onset multiple sclerosis	Unknown	Positive	Negative	Positive
CLP-25	Male	69	Unknown	Unknown	Positive	Positive	Positive
CLP-26	Male	70	Fever, joint pain, EM	Unknown	Positive	Negative	Positive

GROUP 4: CZECH LYME DISEASE NEGATIVE

Sera submitted for Lyme disease testing that are seronegative and have a clinical history that is not consistent with early Lyme disease.

Type: Serum, Retrospective; residual diagnostic samples

Source: Hospital Ceske Budejovice **Table 2.4:** Group 4 sample details

ID	Sex	Age	Additional clinical notes	C6 Lyme Index	C6 interpretation	Virastripe IgM	Virastripe IgG
CLN-1	Male	40	Unknown	Unknown	Negative	N/A	N/A
CLN-2	Female	86	Unknown	Unknown	Negative	N/A	N/A
CLN-3	Female	56	Unknown	Unknown	Negative	N/A	N/A
CLN-4	Male	41	Unknown	Unknown	Negative	N/A	N/A
CLN-5	Female	2	Unknown	Unknown	Negative	N/A	N/A
CLN-6	Male	48	Unknown	Unknown	Negative	N/A	N/A
CLN-7	Male	45	Unknown	Unknown	Negative	N/A	N/A
CLN-8	Female	10	Unknown	Unknown	Negative	N/A	N/A
CLN-9	Male	22	Unknown	Unknown	Negative	N/A	N/A
CLN-10	Male	66	Unknown	Unknown	Negative	N/A	N/A
CLN-11	Male	40	Unknown	Unknown	Negative	N/A	N/A
CLN-12	Female	55	Unknown	Unknown	Negative	N/A	N/A
CLN-13	Male	5	Unknown	Unknown	Negative	N/A	N/A
CLN-14	Male	5	Unknown	Unknown	Negative	N/A	N/A
CLN-15	Female	15	Unknown	Unknown	Negative	N/A	N/A
CLN-16	Female	16	Unknown	Unknown	Negative	N/A	N/A
CLN-17	Female	57	Unknown	Unknown	Negative	N/A	N/A
CLN-18	Male	54	Unknown	Unknown	Negative	N/A	N/A
CLN-19	Female	62	Unknown	Unknown	Negative	N/A	N/A
CLN-20	Female	33	Unknown	Unknown	Negative	N/A	N/A
CLN-21	Female	61	Unknown	Unknown	Negative	N/A	N/A
CLN-22	Female	3	Unknown	Unknown	Negative	N/A	N/A
CLN-23	Female	53	Unknown	Unknown	Negative	N/A	N/A
CLN-24	Female	71	Unknown	Unknown	Negative	N/A	N/A

GROUP 5: LEPTOSPIROSIS POSITIVE
Positive diagnostic samples consistent with acute leptospirosis (seronegative for Lyme disease by STT testing)
Type: Serum
Source: RIPL & BRD, PHE

Table 2.5: Group 5 sample details

ID	Sex	Age	Presenting symptoms	Lepto IgM Index	IgM interpretation	16S PCR	MAT
LEP-1	M	65	Jaundice, malaise, fever	2.511	Positive	Negative	Positive
LEP-2	F	51	Unknown	2.593	Positive	Unknown	Negative
LEP-3	F	20	Fever, headache, flu-like illness	2.668	Positive	Unknown	Negative
LEP-4	F	49	Jaundice, malaise, fever	2.748	Positive	Unknown	Unknown
LEP-5	F	25	Unknown	3.084	Positive	Negative	Unknown
LEP-6	F	22	Athralgia, fever	3.244	Positive	Negative	Unknown
LEP-7	M	29	Fever, rash malaise	3.300	Positive	Negative	Equivocal
LEP-8	M	29	Unknown	3.586	Positive	Negative	Unknown
LEP-9	F	24	Unknown	3.640	Positive	Unknown	Unknown
LEP-10	F	55	Unknown	4.946	Positive	Negative	Unknown
LEP-11	M	30	Fever, headache, flu-like illness	2.655	Positive	Unknown	Unknown
LEP-12	F	61	Jaundice, malaise, fever	2.609	Positive	Unknown	Unknown
LEP-13	M	67	Fever, headache, flu-like illness	2.902	Positive	Unknown	Unknown
LEP-14	M	55	Unknown	3.602	Positive	Unknown	Unknown
LEP-15	F	30	Jaundice, malaise, fever	3.111	Positive	Unknown	Unknown
LEP-16	F	26	Unknown	3.304	Positive	Unknown	Unknown
LEP-17	F	22	Jaundice, malaise, fever	2.512	Positive	Unknown	Unknown
LEP-18	M	61	Jaundice, malaise, fever	2.543	Positive	Unknown	Unknown

GROUP 6: SYPHILIS POSITIVE
Positive diagnostic samples consistent with recent/active syphilis
Type: Serum
Source: BRD, PHE Colindale
Sample IDS: SYP1-SYP10 (No clinical information available)

GROUP 7: NORMAL HEALTHY DONORS 1
Normal healthy donor sera (tested seronegative for Lyme disease by routine testing protocol
Type: Plasma
Source: NHSBT
Sample IDS: H1-H30 (No clinical information available)

GROUP 8: NORMAL HEALTHY DONORS 2
Normal healthy donor sera tested seronegative for Lyme disease by routine testing protocol
Type: Serum
Source: PHE Porton
Sample IDS: HQ1-HQ30 (No clinical information available)

GROUP 9: CZECH LYME DISEASE POSITIVE WHOLE-BLOOD

Whole blood samples from patients with that have a clinical history / symptoms consistent with early Lyme disease. Blood was collected in PAXgene tubes to stabilise RNA.

Type: Whole blood

Source: Hospital Ceske Budejovice

Table 2.6: Group 9 sample details

ID	Patient	Sex	Days since presentation	CLIA IgG	CLIA IgM	WB IgG	WB IgM
1/84	1	Male	84	Negative	Borderline	Negative	Negative
2/92	2	Female	92	Positive	Negative	Positive	Negative
7/17	7	Female	17	Positive	Positive	Positive	Negative
7/30			30	Positive	Negative	Borderline	Negative
7/88			88	Positive	Negative	Positive	Negative
8/1	8	Male	1	Negative	Negative	x	x
8/20			20	Negative	Negative	x	x
9/18	9	Male	18	Negative	Positive	Negative	Positive
9/33			33	Negative	Positive	Negative	Positive
10/1	10	Female	1	Borderline	Positive	Borderline	Positive
11/1	11	Female	1	Positive	Positive	Negative	Positive
11/14			14	Positive	Positive	Negative	Positive
12/1	12	Male	1	Positive	Positive	Borderline	Positive
12/18			18	Positive	Positive	Borderline	Positive

2.2. SAMPLE PREPARATION FOR MASS SPECTROMETRY

2.2.1. Immunodepletion

Serum and plasma samples to be included in mass spectrometry analysis (Materials and Methods 3.5) were immunodepleted using Pierce Top-12 Spin depletion columns (Thermo, Pierce top 12 abundant protein depletion columns, C#85165) according to the manufacturer's standard protocol. **Table 2.7** shows the list of proteins removed. 10 µL of undiluted serum sample was added to the depletion column and incubated at room temperature for an hour with continuous end-over-end mixing. The tubes were then centrifuged at 1000g for 2 min. Following centrifugation the total eluate from the columns was 450μL per sample (depleted sample in 10 mM PBS, 0.15 M NaCl, 0.02% azide, pH 7.4).

List of top 12 proteins removed by Pierce Top 12 depletion columns	
Alpha-1-Acid Glycoprotein	Fibrinogen
Aplha-1-Antitrypsin	Haptoglobin
Alpha-2-Macroglubulin	Immunoglobulin A
Albumin	Immunoglobulin G
Apolipoprotein A-I	Immunoglobulin M
Apolipoprotein A-II	Transferrin

Table 2.7: Top 12 abundant proteins removed using Pierce immunodepletion columns (data according to manufacturer).

Depletion was validated by a bicinconinic acid assay (BCA) to quantify total protein before and after depletion, described below. Further validation involved SDS-PAGE with SYPRO ruby staining and Western blots using antibodies for human serum albumin and C-reactive protein.

2.2.2. BCA procedure

Total protein concentration of serum and plasma samples was determined using BCA, using a commercial kit following the manufacturer's 96 well microplate protocol (Thermo Fisher Scientific, Pierce BCA Kit, C#23225). Briefly, kit standards were created by serially diluting bovine serum albumin [(BSA) (provided in kit)] in distilled water to working concentrations ranging from 25 to 2000 μg/μL. In duplicate, 25 μL of each sample was transferred to a well using a multichannel pipette and mixed with 200μL assay development reagent (provided in kit). Plates were incubated with shaking for 30 min at 37 °C. Clinical samples were diluted 1:50 in PBS to ensure the protein concentration fell within the range of the standard curve. Microtitre plates were read at an absorbance of 570 nm on a TECAN microplate reader. Sample protein concentrations were calculated by reference to the standard curve in MS Excel.

2.3. SDS-PAGE AND WESTERN BLOTTING

2.3.1. SDS-PAGE

Gel preparation: A Bio-Rad Mini-Protein system gel electrophoresis tank was used for SDS-polyacrylamide gel electrophoresis (SDS-PAGE). Resolving gels of appropriate acrylamide concentration and stacking gel (5%) were made as described in **table 2.8**. To cast the gel, immediately after the addition of TEMED to the resolving gel mix, the solution was poured between clean glass plates in a casting frame. A small volume of water-saturated butanol was added to the top of the gel to remove air bubbles and create a level surface for later addition of stacking gel. After

15 min when the gel had set, the water-saturated butanol was removed and the stacking gel solution was added followed by a plastic well comb. Following stacking gel polymerisation (15 min), the comb was removed and the glass plate assembly and gel were placed inside an electrophoresis tank. 1x SDS-PAGE running buffer (25 nM Tris-HCL [pH 8.3], 250 μM glycine, 0.1% SDS) was added to the tank to submerge the gel assembly.

Sample preparation: Sufficient sample buffer was prepared based on the number of unknown samples. 100 μL of beta-mercaptoethanol (B-ME) was added per 900μL of 4x Laemmli buffer (Biorad C#161-0737) to give 20μL per unknown. Samples were diluted 3 parts sample with 1 part 4x Laemmli buffer plus B-ME and then denatured at 90°C for 10 min before being placed on ice.

Running gel: 5 μL of pre-stained protein ladder (New England Biolabs: 11-245kDA protein ladder, C#: P7712S) was loaded to well 1 as a reference for molecular weight. Denatured samples were loaded at a volume of 15 μL or 25 μL depending on the number of wells cast (20 and 10 respectively). The gel was run at 100-150 V, current 200-400 mA for 30 min to 1 hour until optimum separation of proteins (as determined by visual inspection of the pre-stained ladder).

Resolving gel

Gel Concentration	12%	10%	7.5%
Resolution (Protein size)	12-60kDa	20-80kDa	36-94kDa
30% acrylamide	4mL	3.3mL	2.5mL
1.5mM Tris-HCL pH 8.8	2.5mL	2.5mL	2.5mL
ddH20	3.3mL	4mL	4.8mL
10% (w/v) SDS	100uL	100uL	100uL
10% (w/v) APS	100uL	100uL	100uL
TEMED (stock)	10uL	10uL	10uL

Stacking/loading gel

Gel Concentration	5%
30% acrylamide	0.83mL
1.5mM Tris-HCL pH 8.8	0.63mL
ddH20	3.4mL
10% (w/v) SDS	50uL
10% (w/v) APS	50uL
TEMED (stock)	15uL

Table 2.8: Table showing gel compositions (per 2 gels) for SDS-PAGE with recommended resolution (protein size) depending on protein of interest. Stacking/loading gels were consistent regardless of target protein size. APS = Ammonium persulfate.

2.3.2 Western blotting

For Western blotting, a semi-dry blotting method was followed using a Bio-rad semi-dry instrument (Bio-rad, Trans-blot SD Semi-dry transfer cell, C#170-3940). Following SDS-PAGE, gels were removed from the gel tank and glass plates using a plastic scraper. Polyvinylidine fluoride (PVDF) membranes (Merck-Millipore, Immobilon-P PVDF membrane, C#IPVH00010) were briefly submerged in 100% methanol before use. Two pieces of filter paper were soaked in Towbin buffer (25 mM Tris-HCL [pH 8.3], 192 mM glycine, 20% methanol) . The gel, filter paper and

membrane were arranged in the following order in the semi-dry blotting apparatus (bottom to top): filter paper, PVDF membrane, gel, filter paper. Transfers were carried out at 15 V for 1.5 hours. After transfer, the membrane was separated from the other materials and blocked in 5% [w/v] fat free skimmed milk buffer made with Tris-buffered saline [(TBS);50 mM Tris-HCL [pH 8.3], 150 mM NaCl) with 0.5% [w/v] Tween-20 (TBST) for 1 hour at room temperature. The blocking buffer was then discarded and the membrane was washed 3 times, for 5 min each, in TBST on a rocker. Primary antibody was diluted in blocking buffer and 5 mL was added to the membrane, sealed in a plastic pouch and incubated at 4 °C overnight on a rocker. Unbound primary antibody was removed by 3 washes in TBST as previously described. HRP-conjugated secondary antibody was diluted in blocking buffer, added to the membrane and incubated for 1 hour at room temperature on a rocker. Unbound antibody was again removed with 3 washes in TBST. Bound antibody was visualised using enhanced chemiluminescence (ECL). The ECL Western blotting substrate (Bio-rad, Clarity ECL Western substrates, C#1705061) was prepared according to the manufacturer's instructions and 10 mL added to the blot in a plastic dish. After 30 sec of incubation at room temperature, blots were viewed using a Bio-Rad ChemiDoc MP Imaging System. Images were acquired using Image Lab™ Software (Bio-rad, Image Lab, Version: 6.0.1).

Target	Type	Antibody
Rabbit IgG	Secondary for WB	Goat Anti-Rabbit IgG H&L (HRP) (abcam) (ab6721)
Rabbit IgG	Secondary for WB	Goat Anti-Rabbit IgG (abcam) (ab6702)
Lipocalin 2	Primary for WB	Anti-Lipocalin-2 / NGAL antibody (abcam) (ab63929)
Lipocalin 2	Primary for WB	Anti-Lipocalin-2 / NGAL antibody (abcam) (ab41105)
Carbonic anhydrase III	Antibody pair for ELISA	Carbonic Anhydrase III/CA3 Antibody Pair (Novus) (H00000761-AP11)
Serum amyloid A1	Antibody pair for ELISA	Human Serum Amyloid A1 duoset ELISA system (RNDsystems) (DY3019)
C-reactive protein	Primary for WB	Anti-CRP antibody (abcam) (ab227507)
Von Willebrand Factor	Primary for WB	Anti-VWF antibody (abcam) (ab6994)
Tubulin	Primary for WB	Anti-Tubulin antibody (abcam) (ab6160)
Beta-Tubulin	Primary for WB	Anti-beta Tubulin antibody (abcam) (ab6046)

Table 2.9: List of antibodies used in WB and ELISA

2.3.3. SYPRO ruby staining: Whole protein normalisation

Whole protein visualisation of SDS-PAGE gels were performed using SYPRO Ruby protein gel stain (Invitrogen, SYPRO Ruby gel stain, S1200). Following migration, the gel was incubated in fix solution (50% [v/v] methanol / 7% [v/v] acetic acid) for 2 x 30 min at room temperature with shaking. In dark conditions, the gel was incubated with 5 mL of SYPRO gel stain overnight. The stain was then discarded and the gel was gently washed in wash solution (10% [v/v] methanol / 7% [v/v] acetic acid) for 30 min. The gel was rinsed in ddH_{20} for 2 x 5 min before being imaged using a gel imager (Biorad, Chemidoc Touch Imaging System).

2.4. ELISA

2.4.1. Duoset ELISA kits

Quantification of Lipocalin 2 and Serum Amyloid A1 proteins in human sera were performed using DuoSet ELISA systems (R&D Systems, Human Lipocalin 2 Duoset, Human Serum Amyloid A1 Duoset, DLCN20, DY3019-05), according to the manufacturer's protocol. 96-well plates were coated with 100 μL capture antibody diluted to working concentration in PBS and incubated overnight at room temperature. Wells were aspirated and washed with wash buffer (400 μL) for a total of three times. Plates were blocked using 300 μL of kit reagent diluent (BSA based) per well and incubated for 1 hour at room temperature. The aspiration/wash step was repeated as described. 100 μL of sample or standards diluted in reagent diluent to concentration described in the relevant chapter. Plates were covered with an adhesive strip and incubated for 2 hours at room temperature. Following an aspiration/wash step, 100 μL of detection antibody, diluted in reagent diluent, to each well and incubated for 2 hours at room temperature. Following an aspiration/wash step, 100 μL of

Streptavidin-HRP, diluted to working concentration, was added per well. Following an aspiration/wash step, 100 μL of substrate solution was added to each well and incubated for 20 min at room temperature. 50 μL of stop solution (acid) was then added to each well. Optical density was measured immediately, using a microplate reader (Tecan, Infinite 200) set to 450 nm with wavelength correction. Kit standards were used to create a standard curve for the calculation of protein concentration in unknown samples.

2.5. MASS SPECTROMETRY

2.5.1. Tryptic digestion

Serum and plasma samples were depleted using Pierce Top 12 immunodepletion columns as described in section 3.2.1 Samples were then precipitated by addition of ice-chilled trichloroacetic acid (30% [v/v] in acetone) and incubation for three hours at 20 °C. Following precipitation of proteins from solution, samples were centrifuged at 12,000 g for 10 min at room temperature to pellet proteins. Ice-cold acetone was used to wash the protein pellets three times and they were then allowed to air-dry. Pellets were re-suspended in suspension buffer (50 mM ammonium bicarbonate with 0.1% [w/v] RapiGest SF surfactant (Waters, RapiGest, C#87201). Samples were heated to 80 °C on a heatblock. After 10 min, 3 mM dithiothreitol (DTT) was added to reduce samples which were then cooled and alkylated with 9 mM iodoacetamide for 30 min at room temperature. Trypsin (Sigma, SOLu-Trypsin Dimethylated, proteomics grade, C#EMS0005) was added at a ratio of 50:1, sample protein to trypsin by mass, and incubated at 37 °C overnight in the dark. Finally, to remove the

surfactant, the samples were precipitated using 1% (v/v) trifluroacetic acid at 37 °C for two hours and centrifuged for one hour at 12,000g (4°C).

2.5.2. Mass spectrometry analysis: Nano-LC-ESI MS/MS analysis

A nanoscale liquid chromatography coupled to tandem mass spectrometry (nano LC-MS/MS) system was used for protein identification and quantification. Peptides were analysed by nanoflow LC using the Thermo EASY-nLC 1000 liquid chromatograph coupled with a Q-Exactive mass spectrometer (Thermo, Q Exactive hybrid quadrapole-orbitrap, C#IQALAAEGAAPFALGMAZR).

Pre-prepared samples were loaded on an EASY-Spray™ LC column (Length: 50 cm; internal diameter: 75 μm, content: 2 μM C_{18} particles) and then integrated with a nano-electrospray emitter (Thermo, Silica nano-electrospray emmiter, C#ES081). The column was maintained at a temperature of 35 °C throughout operation.

Chromatography: Chromatography was performed using C2v2 buffer A (0.1% formic acid) and B (80% acetonitrile in 0.1% formic acid). Peptides were separated by a linear gradient of 3.8%-50% buffer B over 157 min (flow rate: 300 nl/minute).

Mass spectrometry: The Q-Exactive was operated in a data-dependent mode with survey scans acquired at a resolution of 70,000. The 10 most abundant isotope patterns with charge states of +2, +3 and/or +4 from the initial survey scan were selected with an isolation window of 2.0Th and fragmented by higher energy collisional dissociation with normalized collision energies of 30. The maximum ion infection times for the survey scan and the MS/MS scans were 250 and 100 ms, respectively. The ion target value of the machine was set to 1E6 (survey scans) and 1E4 (MS/MS scans). Repetitive sequencing of peptides was minimized through dynamic exclusion of the sequenced peptides for 20s. RAW files were automatically

transferred for importation into Progenesis software for protein ID and quantification. Mass spectrometry analyses were performed at the University of Liverpool by Stuart Armstrong.

2.5.3. Progenesis protein identification and quantification

RAW files from MS/MS (Thermo RAW) were imported into Progenesis LC-MS software (Nonlinear Dynamics, Progenesis Software, Version: 4.1). Peaks were identified by the software using default settings and filtered (to peaks with charge stage between +2 and +7). Spectral data were converted to MGF files with Progenesis LC-MS and exported for peptide ID using Mascot search-engine software (Matrix Science, Mascot, Version: 2.3.02). Tandem MS data were searched against translated Open Reading Frames (ORFs) from the human genome (Uniprot, 2016_08, 20,367 sequences, 11,398,732 residues) and the *Borrelia burgdorferi* s.l. genome (NCBI Refseq NC_001852, *Borrelia burgdorferi*). The following search parameters were used; Precursor mass tolerance=10, Fragment mass tolerance=0.01Da, Tryptic cleavages permitted=2, Carbamidomethylation=fixed modification, Oxidation=variable modification, Mascot decoy database function=true. Mascot search results were further validated using the embedded semi-supervised machine learning algorithm, Percolator to improve the discrimination between correct and incorrect spectrum identifications. The false discovery rate was <1% with individual Percolator ion scores >13 indicating identity or extensive homology ($p<0.05$). Search results from Mascot were imported into Progenesis LC-MS as XMS files. Relative quantification using (non-conflicting) peptides was calculated by Progenesis. Data was exported in an MS Excel document with normalised abundance data for all proteins successfully identified for each sample.

2.5.4. Data analysis of protein data

Proteins were identified based on detection of at least two unique peptides. The mean normalised abundance for a particular protein was calculated per sample group and fold difference was calculated along with adjusted p-value. Fold changes with p-value <0.05 were considered not significant.

2.6. RNA SEQUENCING AND ANALYSIS

2.6.1. RNA extraction and sample preparation

Whole blood samples were received from the Czech Republic frozen in PAXgene Blood RNA tubes (BD, PAXgene Blood RNA Tubes, 762165). Intracellular blood was isolated and purified using the relevant PAXgene kit (BD, PAXgene Blood RNA kit IVD, 762164) following the manufacturers instructions, as follows: The PAXgene tubes were centrifuged for 10 min at 3000-5000 x g using a swing-out rotor. The supernatant was removed by pipette and 4mL of RNase-free water was added to the pellet. The pellet was dissolved by vortexing and centrifuged again for 10 min at 300-5000 x g. The supernatant was removed by pipette and 350 μL of Buffer BR1 (kit resuspension buffer) was added to the pellet. The sample was pipetted into a new 1.5mL tube and 300 μL of Buffer BR2 (kit binding buffer) and 40 μL of proteinase K were added. The lysate was pipetted into a PAXgene Shredder spin tube and centrifuged for 3 min at 15000 x g. The supernatant was transferred to a new 1.5mL tube leaving the pellet in the original tube. 350 μL of ethanol was added, vortexed and centrifuged briefly (2 sec at 1000 x g) to remove drops in the lid. 700 μL of sample

was pipetted into a PAXgene RNA spin column, placed in a 2mL processing tube and centrifuged for 1 min at 10000 x g. 10 μL of DNase I (RNase-free) was added to the spin column membrane and incubated on the benchtop for 15 min at room temperature. 350 μL of Buffer BR3 (kit wash buffer) was added to the spin column and centrifuged for 1 min at 10000 x g. 500 μL of Buffer BR4 (kit wash buffer II) was added and centrifuged for 1 min at 10000 x g. A further 500 μL of Buffer BR4 was added to the spin column and centrifuged for 3 min at 10000 x g. The column was placed into a new processing tube and 80 μL of Buffer BR5 (kit elution buffer) was added and centrifuged for 1 min at 10000 x g to elute the RNA. The samples were aliquoted for quality and quantification testing. Sample not to be used immediately was stored at -70^{o}C.

2.6.2. Quantifying RNA content and quality

The RNA content of samples was measured using the Qubit RNA system (Thermo, Qubit RNA HS Assay Kit, Q32855). The Qubit working solution was prepared by diluting RNA HS Reagent (kit reagent, 200X concentrate in DMSO) 1:200 in HS buffer (kit reagent). Sufficient working solution to allow 200 μL per sample and standard was prepared. 190 μL of working solution was added to each of the sample tube to be used for standards. 10 μL of each kit standard was added to the relevant tube and vortexed for 2-3 sec. 198 μL of working solution was then added to the tubes for unknowns. 2 μL of each unknown was then added to a tube containing working solution and vortexed for 2-3 sec. All tubes were incubated at room temperature for 2 min. Samples were read using a Qubit 3.0 Fluorometer using the high sensitivity method. The fluorometer automatically calculates sample RNA concentration in ng/mL.

2.6.3. Agarose gel electrophoresis of RNA

Agrose gel electrophoresis with ethidium bromide staining was used to visualise potential RNA degradation. Agarose (AGTC Bioproducts) (1% [w/v]) was melted in 1X TAE buffer (40mM Tris-acetate, 1mM EDTA, pH 8.0) by heating and ethidium bromide was added to molten agarose to give a final concentration of 0.5 μg/mL. The molten agarose was added to a gel-casting tray containing a 20 sample comb, and allowed to set. The set gel was transferred to a gel tank and submerged with sufficient 1 x TBE buffer to cover the gel. 0.5 μg of RNA sample in 10μL RNase free water, together with 2 μL of loading dye (Thermo, RNA gel loading dye (2X), R0641) was loaded into each well. The gel was run at a voltage of 2.5 V/cm and samples were migrated until the gel-loading buffer had reached ¾ of the way down the gel. An ultraviolet transilluminator was then used to visualise the RNA bands. Any smearing of RNA bands would indicate a degraded sample.

2.6.4. Illumina RNA-seq

Samples were taken on ice to the Centre for Genomic Research at the University of Liverpool for sequencing. To include only mRNA data, 3'polyadenylated poly(A) selection was used and ribosomal RNA (rRNA) was depleted. Poly(A) selection was performed using poly(T) oligomers covalently bound to a magnetic bead substrate. RNA was reverse transcribed to cDNA for amplification. Sequencing data was provided as "trimmed read data", with poor quality bases at the ends of reads removed.

2.6.5. RNAseq: Data analysis

Downstream processing from trimmed reads provided by RNA-seq to Differential Gene Analysis (DGE) were performed using the Galaxy online platform (Galaxy, Version: 19.09.rc1) using the software workflow shown in **figure 2.1.**

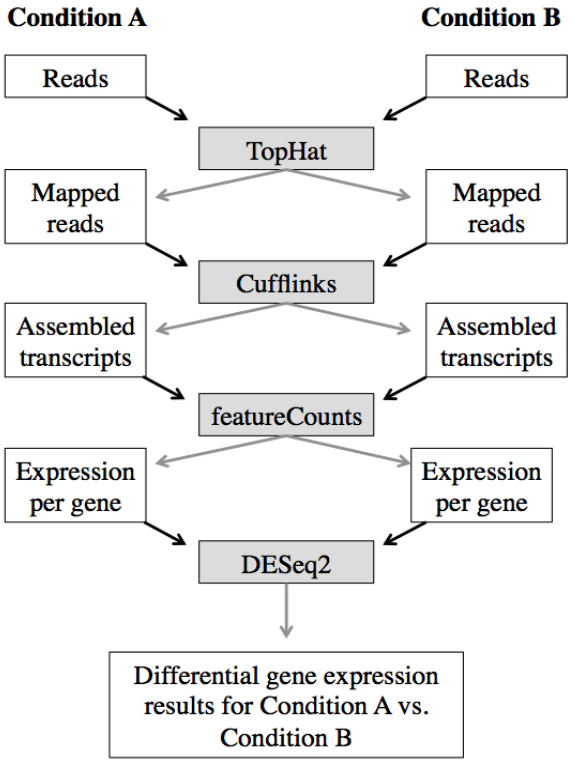


Figure 2.1: Diagram of software workflow used for processing trimmed read data from RNA-seq.

2.6.5.1. Genome mapping

TopHat software was used to map RNA-seq reads to a human reference genome (*Homo sapiens* (b38, hg38)). TopHat uses Bowtie to map RNA-seqs to the reference genome and then analyses the mapping results to identify splice junctions between exons. All default TopHat settings were used. Files were output as accepted hits in BAM format. Cufflinks (University of Washington, Cufflinks,Version 2.2.1) software

was used to assemble transcripts and estimate their abundance. The BAM files was input and Cufflinks then estimated the relative abundance of transcripts based on the number of reads mapped to each, taking into account biases potentially introduced in library preparation protocols. To summarise this data, featureCounts (Unknown, featureCounts, Version: 1.6.4) software was used. The featureCounts software outputs data in a tab-separated file for each sample, listing all human genes and the number of reads mapping to that gene.

2.6.5.2. Differential gene expression: DESeq2

The DESeq2 package (Love *et al*. 2014, DEseq2, Version: 1.25.10) provides methods to test for differential expression by use of negative binomial generalised linear models. featureCount files were used as input with one file per sequenced sample. Samples were split into "exposure" groups to allow DGE comparison. Pre-filtering of low count genes was performed before running the DESeq2 functions to reduce the amount of memory required and to increase the transformation and testing functions. DESeq calculates a Log [2] fold change for each gene input, using one exposure group as a reference. The software also calculates a Wald test *p* value for each gene.

2.6.6. Immquant cell typing

ImmQuant cell typing was used to estimate the abundance of specific immune cells in blood samples using computational deconvolution by the ImmQuant software (omicX, ImmQuant, Version: N/A). Patient data was input into the software as relative expression profiles from DESeq2 (calculated as the transcription fold-change). The software contains a reference dataset consisting of transcriptional signature of immune-cell subsets and a list of informative marker genes. The software then

decomposes the relative expression profile into cell-type differences between the two samples or sample sets.

2.6.7. Quantitative real-time RT-PCR

Quantitative RT-PCR reactions were performed using the Qiagen OneStep RT-PCR platform (Qiagen, OneStep RT-PCR, 210212). A reaction master mix was prepared as shown in **table 2.10.** A negative control (master mix without template RNA) was included in every experiment. Master mix was mixed thoroughly and dispensed into PCR tubes (total 50 μL per reaction). Template RNA was input at 20 ng per reaction (RNA content of samples was measured by Qubit as described above (2.6.2). The PCR machine (Bio-rad, CFX96) was set to the cycling times and temperatures as described in **table 2.11.** Primers used for specific genes are detailed in **table 2.12.**

Component	Volume/reaction	Final concentration
Master mix		
RNase-free water	Variable	-
5 X QIAGEN Onestep RT-PCR Buffer	10.0 μL	1X
dNTP Mix (containing 10 mM of each dNTP)	2.0 μL	400 μM of each dNTP
Primer A	Variable	0.6 μM
Primer B	Variable	0.6 μM
QIAGEN OneStep RT-PCR Enzyme mix	2.0 μL	-
Template RNA	Variable	5-10 units/reaction
Total volume	**50.0 μL**	-

Table 2.10: PCR reaction master mix recipe and final composition

Step	Temperature	Time
Reverse transcription	50oC	30 min
PCR activation	95oC	15 min
3-step cycling		
Denaturation	94oC	30 sec
Annealing	55oC	30 sec
Extension	72oC	1 min
Number of cycles	25-40 cycles	
Final extension	72oC	10 min

Table 2.11: Thermal cycling times and temperatures for RT-PCR

Gene	Forward primer	Tm	Reverse primer	Tm
AKT i	ACAACCAGGACCATGAGAAG	51.3	ATGAGGGGATGGAGGTGTAG	51.9
AKT ii	TCAAGAAGGACCCCAAGCAG	53.7	GGTAGCAAAGCACCACAGAG	53.1
eIF2a i	CAATCTCCCAGGCTCAAAC	52.6	CCCCCATCTCTACAAAACATAC	45.5
eIF2a ii	CTCCCAGGCTCAAACAATAC	50.4	CCCCCATCTCTACAAAACATAC	50.4
eIF4G i	CTCCCTCAACTCCTTCTCTC	50.6	AAGCACCAATCCCTACCAC	51.2
eIF4G ii	TCCCTCAACTCCTTCTCTCC	51.2	AAGCACCAATCCTACC	51.2
XIAP i	ATTCTTCTGCCTCAGCCTCC	53.6	AGTTCTCCTGCTTCAGCCTC	53.5
XIAP ii	GCAACATTCAACTCCAGCC	51.5	GCCTGTAATCCCAGCTACTC	51.7
XIAP iii	GCAACATTCAACTCCAGCC	51.5	CCCGTCTCCACTGAAAATAC	50.2
eIF5 i	CATTCATTCTCAAAAACCCACC	50.9	CCCAGTCTCCCTTACATTCC	50.4
eIF5 ii	GAAAAACAGAAAGGGCAAAGAC	51.0	CATTTCCACCCACCTCAAC	51.1
eIF5 iii	TATCCAGCAGTGAGACACC	50.4	TTTCAACATTTCCACCCACC	50.6

Table 2.12: Primer list for RT-PCR

2.7. LYME DISEASE DIAGNOSTIC TESTING (STT)

2.7.1. STT Protocol

Serum samples from NHD, leptospirosis, syphilis and CFS cases were tested to determine any seroreactivity that might suggest previous exposure to *Borrelia burgdorferi* spp. pathogens. The work was undertaken under the guidance of laboratory staff at the Porton. PHE Rare and Imported Pathogens Laboratory (RIPL),

provides the national Lyme disease testing service for England and Wales and uses well validated CE-marked assays that conform to ISO 15189:2012.

The standard two-tiered (STT) algorithm for Lyme serodiagnosis was followed, using a C6 screening ELISA followed by (in C6 positive/indeterminate samples) a Virachip IgG and IgM confirmatory immunoblot assay. Routinely, diagnosis of Lyme disease must be made based on patient history (including tick-bite), clinical signs and symptoms (including erythema migrans) and other laboratory testing data in addition to the serological based methods detailed here.

2.7.2. C6 ELISA

All sera/plasma were first tested using a commercially available C6 screening ELISA kit (Immunetics, C6 Lyme ELISA kit, C#DK-E352-096) for the presumptive detection of IgG and/or IgM antibodies against *B. burgdoerferi* s.l. (optimised for circulating European species). The antigen used in the assay is a synthetic C6 peptide derived from the VlsE protein of *B. burgdorferi* s.l.

The following procedure was followed using an automated ELISA system (Dynex Technologies, DS2 ELISA Processing System, C#62000). 100 μL of diluted positive control (10 μL in 90 μL kit diluent) was added to one microwell and 100 μL of kit negative control to the adjacent cell. 100 μL diluted calibrator was added to three additional microwells. 100 μL of each diluted patient sample[1:100 ddH_{20}] was then added to available microwells and incubated at room temperature for 30 min. Wells were aspirated three times using Wash Buffer (supplied). 100 μL of kit conjugate was added into each well and incubated for 20 min at room temperature. Wells were again aspirated and 4 washes performed as above. 100 μL of tetramethylbenzidine (TMB).

ELISA substrate was added into each well and incubated for 4 min at room temperature. 100 μL of kit stop solution into each well and the contents gently mixed. Plates were read immediately after addition of stop solution. Absorbance was read at 450nm with a reference filter of 650nm using an ELISA plate reader (Dynex, MRX Revelation, C#N/A). The following quality control parameters were followed:

- Negative control value (A_{450}) <0.18
- Calibrator values (A_{450}) 0.400-2.00
- Positive control value (A_{450}) >1.2

Any control value being out of range resulted in the plate being discounted and the assay repeated. The assay cutoff value was determined by dividing the mean calibrator A_{450} values by 2.150 (correction coefficient). The Lyme Index value (LI) was calculated for each patient sample by dividing the A_{450} value for the sample by the cutoff value. **Table 2.13** below shows the interpretation guidance as described in the kit manual. Patient samples that tested negative by C6 ELISA were considered seronegative for Lyme disease. Samples found to be positive or equivocal went forwards for confirmatory immunoblot analysis by ViraChip.

Lyme Index value	Interpretation
$\leq$ **0.90**	Negative result. No antibody to *B. burgdorferi* detected in the present assay. This result does not exclude the possibility of *B. burgdorferi* infection, and where early Lyme disease is suspected, a second sample should be drawn 2 4 weeks later and re-tested.
0.91-1.09	Equivocal result. The imprecision inherent in any method implies a lower degree of confidence in the interpretation of samples with A450 values very close to the calculated cutoff value. For this reason an equivocal category has been designated. Equivocal samples should be tested with a supplemental assay such as a standardized Western Blot test in accordance with CDC/ASTPHLD recommendations.
$\geq$ **1.10**	Positive result. Antibody to *B. burgdorferi* detected in the present assay. All positive results should be supplemented by re-testing the corresponding serum samples on a standardised Western Blot test in accordance with CDC/ASTPHLD recommendations.

Table 2.13: Interpretation of sample Lyme Index values as described in Immunetics C6 ELISA kit manual.

2.7.3. ViraChip Borrelia Immunoblot: IgG and IgM

Confirmatory testing used chip-based immunoblots for the qualitative detection of IgG and IgM antibodies against specific *Borrelia* species antigens. Separate kits were used for detection of *Borrelia* specific IgM (Viramed; Borrelia ViraChip IgM; C# V-BSCMOK) and IgG (Viramed; Borrelia ViraChip IgG; C# V-BSCGOK) antibodies. Both kits use native antigens from *Borrelia afzelli* and *Borrelia burgdorferi s.s.* together with recombinant VlsE, applied (in triplicate) at defined positions in each well of a 96-well microtitre plate. Using a TECAN automated ELISA instrument, the following protocol was followed for both IgG and IgM plates. 300 μL of kit diluent was added to each well. 100 μL of diluted kit positive control, 100 μL of kit negative control, 100 μL of in-house positive control (PHE) and 100 μL of in-house negative control (PHE) was added to the appropriate wells. 100 μL of diluted patient serum

was then added to the remaining available wells. Plates were incubated for 30 min at room temperature. Wells were then aspirated and washed 3 times in kit wash buffer working dilution (300 μL). 100 μL of kit conjugate was added to the wells followed by another 30 minute incubation at room temperature. Cells were again aspirated and washed as before. 300 μL of deionised water was added to each well and the plate incubated for 5 min at room temperature. The liquid was aspirated and 100 μL of Chromogen substrate solution added to each well and the plate incubated for 15 min at room temperature. The reaction was halted by aspiration of all liquid from wells. A final wash was performed 3 times and the plate was left to air dry. Interpretation of the assay involved measuring of spot intensities of the microarrays using ViraChip software. For a test run to be valid, the serum positive control, conjugate control and calibrator controls must all be detectable. Tests in which the negative controls were visible were not valid. Automated spot assessments were manually confirmed by eye using the ViraChip software. **Table 2.14** describes the interpretation of results, as described in the ViraChip IgG and IgM manual.

IgG ViraChip interpretation

Identified spot triplet	Result	Interpretation
At least two distinct spot triplets out of: p83, p58, p43, p39, p30, OspC, p21, Osp17 or DbpA, p14, VlsE	Positive	Specific IgG antibodies against Borrelia species detectable. An infection with Borrelia species is probable.
One distinct VlsE spot triplet	Equivocal	Specific IgG antibodies against VlsE detectable. An infection with Borrelia species is possible. If a recent infection is suspected, check additionally for IgM antibodies. If possible test a second sample for IgM and IgG specific antibodies after 2-3 weeks.
One or no distinct spot triplet (Exception: singular VlsE band)	Negative	No specific IgG antibodies against Borrelia species detectable. If an infection is suspected, check a second sample for IgM and IgG specific antibodies after 2-3 weeks.

IgM ViraChip interpretation

Identified spot triplet	Result	Interpretation
At least one distinct spot triplets out of: p41, p39, OspC, Osp17 or VlsE	Positive	Specific IgM antibodies against Borrelia species detectable. An infection with Borrelia species is probable.
No or only weak spot triplets	Negative	No specific IgM antibodies against Borrelia species detectable. If an infection is suspected, check a second sample for IgM and IgG specific antibodies after 2-3 weeks.

Table 2.14: ViraChip result interpretation as detailed in the kit manual.

2.8. SEROLOGY LITERATURE REVIEW PARAMETERS

2.8.1. Literature search For table 1.1-1.4

Primary research questions of review: Sensitivity and specificity of serological tests for the conditions early Lyme (EM), Lyme arthritis, ACA and neuroborreliosis. Studies focusing on putative high-risk groups including veterinarians and forestry workers were excluded. Serology-based tests associated with STTT were included including commercial and in-house. ELISA, IB and two-tier algorithm tests were identified. Quality score was based on a version of QUADAS-2. Four domains were

considered for quality score including patients cohort selection, index test and risk of bias. **Table 2.15** shows search terms used across several databases including EMBASE, PubMed and Web of Science. Irrelevant studies were excluded including those on other diseases, studies in (non-human) animals or where the full study text was not accessible.

Lyme borreliosis	EIA.ti,ab,ot.
Lyme disease	immunosorbent.ti,ab,ot.
immunoglobin*.ti,ab,ot.	Enzygnost.ti,ab,ot
immunoblot*.ti,ab,ot.	VIDAS
'Western blot'.ti,ab,ot.	Serion.ti,ab
exp serology	recomline.ti,ab,ot
euroimmunoblot	ELISPOT.ti,ab,o
mikrogen$.ti,ab,ot.	virotech
tick borne disease/	erythema migrans
Neuroberreliosis.ti,ab,ot	Acrodermatitis chronica atrophicans/
meningoradiculitis	Two-tier
erythema chronicum migrans/	borrelia.ti,ab,ot
burgdorferi.ti,ab,ot	lymphocyte transformation test

Table 2.15: Search terms for literature search regarding serological testing for Lyme disease

CHAPTER 3: MASS SPECTROMETRY BASED DISCOVERY OF SERUM BIOMARKERS OF SPIROCHAETAL INFECTIONS

3.1. INTRODUCTION

Current Lyme disease diagnostic methods used in the UK, the STT method, are based on the detection of antibodies against the *Borrelia* pathogen in the patient's blood. In early stages of Lyme disease, levels of *Borrelia*-specific antibodies are low as the host mounts an immune response to infection. During this time, serology based tests have lower sensitivity as they rely on the detection of host antibodies to several bacterial proteins. Studies vary in their estimation of the length of time before seroconversion, when a patient with Lyme disease has sufficient levels of antibody to test positive by STT, but it is likely to be at least two weeks or longer (217)

Blood is a commonly used tissue for diagnostic purposes, due to the ease and safety of sampling. Blood samples collected by phlebotomy remain the most common type of biological sample drawn for analysis, supporting patient diagnosis and monitoring (218). Whole blood can be processed into serum or plasma fractions, depending on the intended use of the sample. Plasma, which constitutes around 55% of whole blood, contains fibrinogen and other secondary clotting factors that contribute to the normal clotting of blood. In contrast, serum is prepared by allowing whole blood to form clots by the natural conversion of fibrinogen to fibrin. The clot is then removed by centrifugation leaving serum, the liquid portion of blood minus clotting factors and cells.

Measuring the abundance of resident blood proteins, together with proteins from tissues that can enter the bloodstream, can allow the detection of putative biomarkers associated with a disease state and for better understanding of the host-response to disease. Mass spectrometry analyses can be used to quantify proteins in a complex biological sample by detection of peptide subunits. Improvements in coverage and depth of mass spectrometry methods have allowed better resolution of the blood proteome in individual patients, leading to studies with the aim of identifying biomarkers for certain diseases, disease states and sequelae, as well as other aspects such as host response to treatment including antibiotic therapy. A biomarker can be defined as any measurable characteristic that can be quantified as an indicator of pathological processes, a host-response to a treatment or exposure, or simply of normal biological processes. In this chapter, biomarkers refer to serum proteins identified by mass spectrometry analysis that are differentially abundant in different diseases and disease states.

This chapter reports the findings of two separate label-free quantitative mass spectrometry studies performed on Lyme disease serum samples and related-disease control and normal healthy donor samples with the aim of characterising the serum proteome of early Lyme disease patients and identifying serum biomarkers for further evaluation as novel biomarkers of early Lyme disease.

3.1.1. Hypothesis

The serum proteome of patients with Lyme disease, as diagnosed based on clinical criteria and positive serology using a standard diagnostic protocol, is distinct from that of cases that test negative for Lyme disease and from disease and healthy control groups. Mass spectrometry analyses can be used to identify proteins that are found consistently at different abundance between groups and these proteins can be taken forward for further research as potential biomarkers of early Lyme disease.

3.2. MATERIALS AND METHODS

Full details of the Materials and Methods used in the chapter are provided in Chapter 2. **Table 3.1** below shows the samples including in mass spectrometry runs 1 and 2 with details of their inclusion criteria and source. Details of individual patient clinical history, diagnostic testing results and other information (where available) are included in Chapter 2: Materials and Methods (2.1.1).

Sample set	Number	Criteria for Inclusion	Sample type	Provider
UK Lyme positive (IDs: LP1-LP26)	26	Sera submitted for Lyme disease testing that are seropositive and have a clinical history / symptoms consistent with early Lyme disease.	Serum - Retrospective; residual diagnostic samples	RIPL, PHE Porton
UK Lyme negative (IDs: LN1-LN17)	17	Sera submitted for Lyme disease testing that are seronegative and have a clinical history that is not consistent with early Lyme disease.	Serum - Retrospective; residual diagnostic samples	RIPL, PHE Porton
Czech Lyme positive (IDs: CLP1-CLP26)	26	Sera submitted for Lyme disease testing that are seropositive and have a clinical history / symptoms consistent with early Lyme disease.	Serum - Retrospective; residual diagnostic samples	Hospital Ceske Budejovice, Czech Republic
Czech Lyme negative (IDs: CLN1-CLN24)	24	Sera submitted for Lyme disease testing that are seronegative and have a clinical history that is not consistent with early Lyme disease.	Serum - Retrospective; residual diagnostic samples	Hospital Ceske Budejovice, Czech Republic
Syphilis positive (IDs: SYP1-SYP10)	10	Positive diagnostic samples consistent with recent/active syphilis (seronegative for Lyme disease by STT testing)	Serum - Retrospective; residual sera provided by NHSBT	BRD, PHE Colindale
Leptospirosis positive (IDs: LEP1-LEP18)	18	Positive diagnostic samples consistent with acute leptospirosis (seronegative for Lyme disease by STT testing)	Serum - Retrospective; residual sera provided by NHSBT	RIPL, PHE Porton. BRD, PHE Colindale
Chronic fatigue syndrome (CFS)(IDs: CFS1-CFS15)	15	Residual sera from previous CFS research study (FATIGUE study)	Previous research study; patient consent for sample storage	UoL
Normal Healthy Donors 1 (IDs: H1-H30)	30	Normal healthy donor sera (tested seronegative for Lyme disease by routine testing protocol)	Plasma - Retrospective; residual sera provided by NHSBT	NHSBT Services
Normal Healthy Donors 2 (IDs: HQ1-HQ30)	30	Normal healthy donor sera (tested seronegative for Lyme disease by routine testing protocol)	Serum - NHD serum aliquots	PHE Porton

Abbreviations: RIPL Rare and Imported Pathogens Laboratory, PHE Porton; BRD Bacteriology Reference Department, PHE Colindale; NHSBT National Health Service Blood and Transfusion; UoL - University of Liverpool.

Table 3.1: Details of serum and plasma samples used during mass spectrometry run 1 and run 2.

3.3. RESULTS

3.3.1. Study sample set

A total of 50 human serum and plasma samples were compared by quantitative label-free mass spectrometry. **Table 3.1** shows the sample sets within the study, including sample type and inclusion criteria. All sample sets comprised serum samples apart from those from patients diagnosed with chronic fatigue syndrome (n=5) and normal healthy donors (n=5), which were plasma. Serum and plasma are both derived from whole blood, but while plasma contains clotting factors including fibrinogen serum samples are centrifuged to remove these leaving only the fluid fraction. The importance of this distinction in mass spectrometry analysis is discussed later in this chapter.

3.3.2. Sample preparation for mass spectrometry

The dynamic range of protein abundance within blood spans several orders of magnitude. Albumin, the most abundant blood protein, constitutes around 55% of the entire blood proteome, over 10 orders of magnitude in concentration greater than the rarest proteins measured clinically. Globulins and fibrinogen constitute up to 38% and 7% of whole-blood proteins respectively. Together, high-abundance proteins represent more than 99% of the total bulk mass of protein content. Removal of abundant proteins by immuno-affinity baits can significantly reduce the complexity and dynamic range of serum and plasma samples, allowing better resolution of low-abundance proteins that may be of clinical interest. Samples were received in either serum or plasma form and underwent protein depletion using Pierce Top 12 Abundant Protein Depletion Spin Columns. The proteins depleted by this method are listed in Materials and Methods 2.2.1. To quantify the amount of protein removed by the

depletion columns, total protein was measured before and following immuno-depletion using a bicinchoninic acid assay (BCA). **Table 3.2** shows this data. The average decrease in total protein following immuno-depletion was 77.43% (75.38-77.47%). This is in line with the depletion efficiency stated by the manufacturer, >99% of albumin and >90% of globulins. To visualise the removal of highly-abundant proteins, SYPRO ruby was used to stain total protein in samples before and after depletion. **Figure 3.1** shows a Western blot of serum samples (Lyme disease seropositive and leptospirosis positive) before and after immuno-depletion, stained by SYPRO ruby. The oversaturated band at around 66.5kDa, assumed to be albumin, can be seen to be reduced following depletion. Better resolution of distinct bands can be seen across a range of molecular weights in the post-depletion samples.

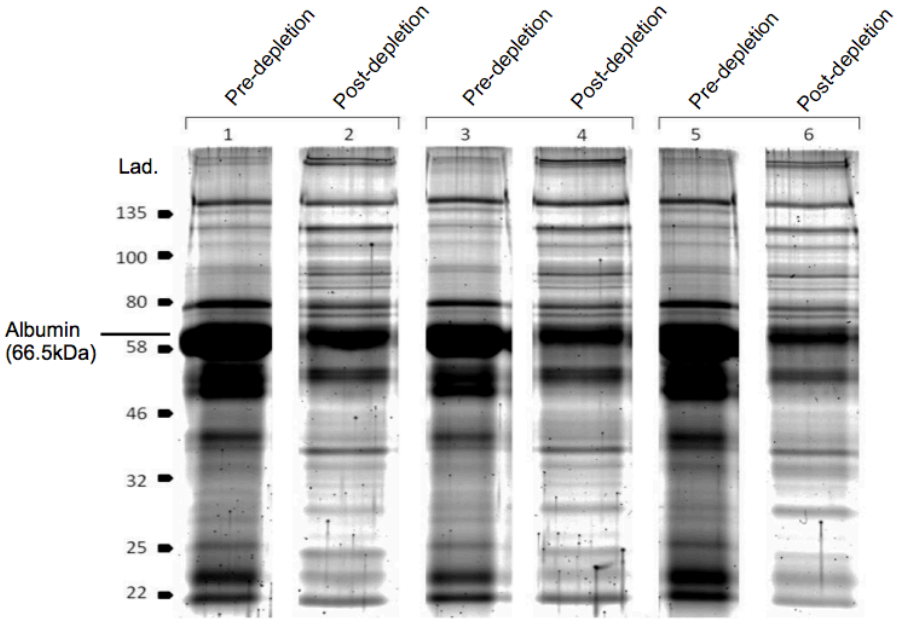


Figure 3.1: Western blot showing serum samples before and after depletion by immuno-affinity columns. Total protein is stained using SYPRO ruby. The oversaturate band at the molecular weight of albumin can be seen to be reduced following depletion. Lanes 1-2: Lyme disease positive sample (LP 3), Lanes 3-4: Leptopsirosis positive sample (LEP 4), Lanes 5-6: Normal healthy control sample (H1)

ID	Pre-depl. protein conc. ($\mu g/\mu L$)	Post-depl. protein conc. ($\mu g/uL$)	Decrease (%)	ID	Pre-depl. protein conc. ($\mu g/\mu L$)	Post-depl. protein conc. ($\mu g/\mu L$)	Decrease (%)
LP 1	110.63	18.73	83.07	**LN 13**	143.12	25.62	82.10
LP 2	108.02	20.54	80.99	**LN 14**	118.68	45.09	62.01
LP 3	111.43	21.00	81.15	**LN 15**	144.44	26.10	81.93
LP 4	104.57	24.01	77.04	**LEP 1**	87.66	24.34	72.24
LP 5	111.25	21.75	80.45	**LEP 2**	99.78	33.12	66.81
LP 6	110.62	27.82	74.85	**LEP 3**	95.69	29.58	69.08
LP 7	105.31	23.95	77.26	**LEP 4**	118.79	25.46	78.57
LP 8	106.41	24.07	77.38	**LEP 5**	89.14	24.27	72.78
LP 9	159.95	29.11	81.80	**SYP 1**	114.51	21.48	81.25
LP 10	129.64	26.00	79.95	**SYP 2**	115.98	24.29	79.06
LP 11	124.12	26.36	78.77	**SYP 3**	119.32	22.22	81.37
LP 12	131.25	27.73	78.87	**SYP 4**	118.54	30.08	74.62
LP 13	94.49	23.19	75.45	**SYP 5**	118.36	23.50	80.15
LP 14	113.12	22.08	80.48	**H 1**	140.96	39.67	71.86
LP 15	118.81	21.53	81.88	**H 2**	109.58	42.71	61.02
LN 1	94.29	18.49	80.39	**H 3**	82.17	22.40	72.75
LN 2	101.79	23.49	76.92	**H 4**	130.84	26.26	79.93
LN 3	106.38	20.19	81.02	**H 5**	206.58	24.53	88.12
LN 4	115.29	26.88	76.68	**H 6**	167.88	28.90	82.78
LN 5	106.01	25.00	76.42	**H 7**	240.96	29.13	87.91
LN 6	133.30	39.75	70.18	**H 8**	183.96	24.31	86.79
LN 7	109.54	30.12	72.50	**CFS 1**	204.68	30.24	85.22
LN 8	114.87	46.01	59.94	**CFS 2**	187.96	24.32	87.06
LN 9	113.96	28.57	74.93	**CFS 3**	167.86	26.35	84.31
LN 10	118.23	62.79	46.90	**CFS 4**	201.03	29.99	85.08
LN 11	139.86	13.78	90.14	**CFS 5**	111.12	26.23	76.39
LN 12	128.42	33.67	73.78	**CFS 6**	119.68	22.65	81.07

Pre-depl. mean ($\mu g/\mu L$)	**Post-depl. mean** ($\mu g/uL$)	**Decrease** (%)
127.05±8.79	27.39±2.12	77.43±2.04

Table 3.2: Total protein depletion by Pierce Top 12 spin columns, as measured by BCA. ID represents individual sample ID. Protein concentration is shown before and following depletion, along with percentage reduction. The mean concentration pre- and post-depletion, with mean percentage reduction are shown at the bottom of the table.

3.3.3. Label-free mass spectrometry run 1

3.3.3.1. Overview

To identify differences in serum proteome between groups a label-free quantitative mass spectrometry study was undertaken. Mass spectrometry run 1 compared the following sample groups: UK Lyme disease seropositive (n=15), UK Lyme disease seronegative (n=15), Syphilis positive (n=5), Leptospirosis positive (n=5), Normal/health donors 1 (Plasma) (n=15), and Chronic Fatigue Syndrome diagnosed (Plasma) (n=15).

A total of 296 proteins were successfully identified and their abundance quantified in all study groups included in mass spectrometry run 1. Mass spectrometry data was normalised between sample groups using Progenesis software as described in Materials and Methods 2.5.3. All samples were plotted using principal component analysis (PCA), an orthogonal transformation that allows complex mass spectrometry data to be shown on a 2D graph The PCA allows a visualisation of similarity between sample groups at the proteome level. **Figure 3.2** shows clear separation of leptospirosis positive samples from the other sample groups, reflecting a distinctive proteome in this sample set. Syphilis positive samples can be seen to group close together, reflecting a relatively consistent proteome across samples. Overlap between syphilis and the Lyme disease groups (seropositive and seronegative) can be seen. Lyme disease seropositive and Seronegative groups show near complete overlap, with Lyme disease seropositive samples sowing closer grouping together. This indicates Lyme disease seropositive and seronegative sera share a broadly similar proteome that is consistent across samples.

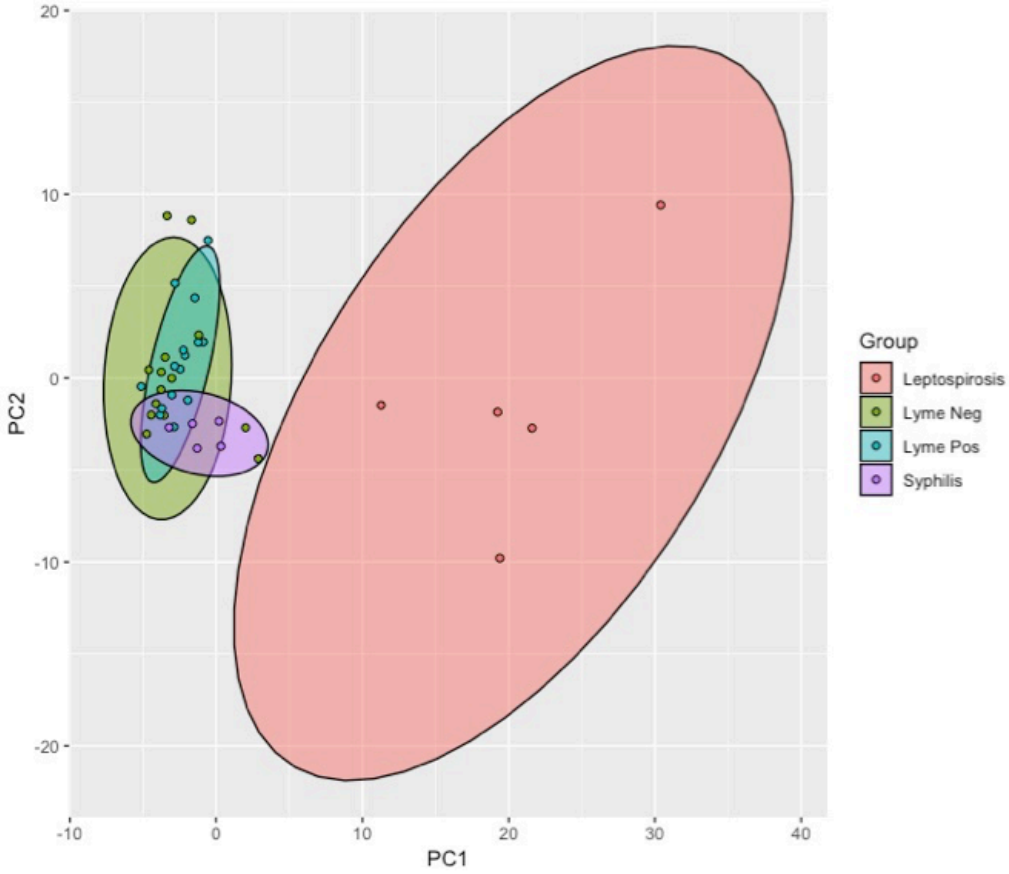


Figure 3.2: Principal Component Analysis (PCA) plot showing normalised data from mass spectrometry run 1. Each point represents a single sample, with colour overlay showing sample groups.

3.3.3.2. Mass spectrometry run 1: Lyme disease seropositive vs. seronegative

Of the proteins identified between Lyme disease seropositive and seronegative groups, 12 proteins were found consistently at significantly different abundances between groups ($p<0.05$). Of these, 9 were found to be increased in abundance in seropositive and 3 decreased in abundance in seropositive. **Figure 3.3** shows a volcano plot comparing the abundance of individual proteins in the Lyme disease seropositive group compared to the seronegative group. A cut-off value for Log[2] Fold Change of either >1 or <-1 was applied with significantly changed proteins shown in red. The protein Lipocalin-2 (LCN2) shows the highest significant increase in abundance in the Lyme disease seropositive group ($p=0.33$, Log[2] Fold Change =2.323). Lipocalin-2

was taken forward for further analysis in chapter 5. **Table 3.3** shows the other significantly altered proteins between these 2 groups. The proteins identified were found to have a range of biological functions.

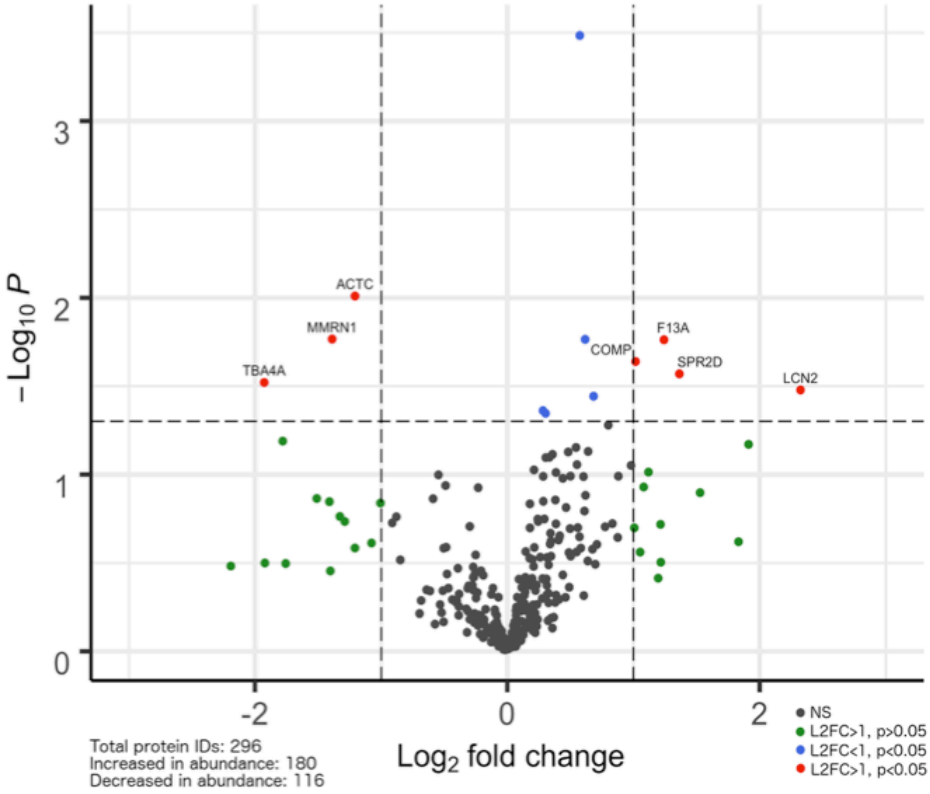


Figure 3.3: Volcano plot using mass spectrometry run 1 data comparing Lyme disease seropositive group to Seronegative. Proteins on the right of the graph are increased in abundance in Lyme disease seropositive, with proteins decreased shown on the left. A p-value of 0.05 (Log[10]p=1.4) was used as a significance cut-off along with a value of 1 for fold change (Log[2]). Proteins with values above these cut-offs are shown in red. HGNC shorthand gene names are given, full names are given in subsequent tables.

UniProt ID	Gene	Protein Name	GO - Biological Function	Log^2FoldChange	*p*-value
Proteins increased in abundance in Lyme seropositive					
P80188	NGAL	Lipocalin 2	Antimicrobial humoral response	2.323	0.033
P22532	SPR2D	Small proline-rich protein 2D	Cornification, epidermis development	1.363	0.027
P00488	F13A	Coagulation factor XIII A chain	Blood coagulation	1.241	0.017
P49747	COMP	Cartilage oligomeric matrix protein	Cornification, Animal organ morphogenesis	1.017	0.023
P54108	CRIS3	Cysteine-rich secretory protein 3	Defence response	0.684	0.036
P47929	LEG7	Galectin-7	Apoptotic process	0.617	0.017
O00391	QSOX1	Sulfhydryl oxidase 1	Cell redox homeostasis	0.575	0.0003
P08185	CBG	Corticosteroid-binding globulin	Glucocorticoid metabolic process	0.304	0.045
P13671	CO6	Complement component C6	Complement activation	0.283	0.044
Proteins decreased in abundance in Lyme seropositive sera					
P68366	TBA4A	Tubulin alpha-4A chain	Ciliary basal body-plasma membrane docking	-1.926	0.03
Q13201	MMRN1	Multimerin-1	Blood coagulation	-1.389	0.017
P68032	ACTC	Actin, alpha cardiac muscle 1	Actin filament-based movement	-1.208	0.009

*(adjusted $p<0.05$)

T able 3.3: Table showing the 9 proteins found to be significantly changed in abundance when comparing Lyme disease seropositive and seronegative groups.

3.3.3.3. Mass spectrometry run 1: Lyme disease seropositive vs. other sera

To identify proteins that are differentially abundant between the Lyme disease seropositive group and all other sera groups, normalised abundance data was compared between the subsets. Including data from related-disease controls allowed identification of proteins that were differentially abundant in Lyme disease patients but not in other disease. 22 proteins were found to be significantly changed. Of these, 19 were found to be increased in abundance in Lyme disease seropositive samples and 3 were decreased in abundance. While a greater number of proteins were significantly changed than when the Lyme disease seropositive and seronegative groups were compared, the average fold change differences are smaller, with only 2 proteins showing a significant increase or decrease with a fold change of >1 or <-1. Coagulation factor XIII alpha chain (Factor XIII) was found to be increased in Lyme disease seropositive ($p=0.02$, Log[2]Fold Change =1.241). Tubulin alpha-4A chain was found to be decreased in Lyme disease seropositive ($p=0.06$, Log[2]Fold Change = -1.208). **Table 3.4** shows all proteins found to be at significantly different abundances between groups. Several of the proteins increased in Lyme disease seropositive are associated with blood coagulation.

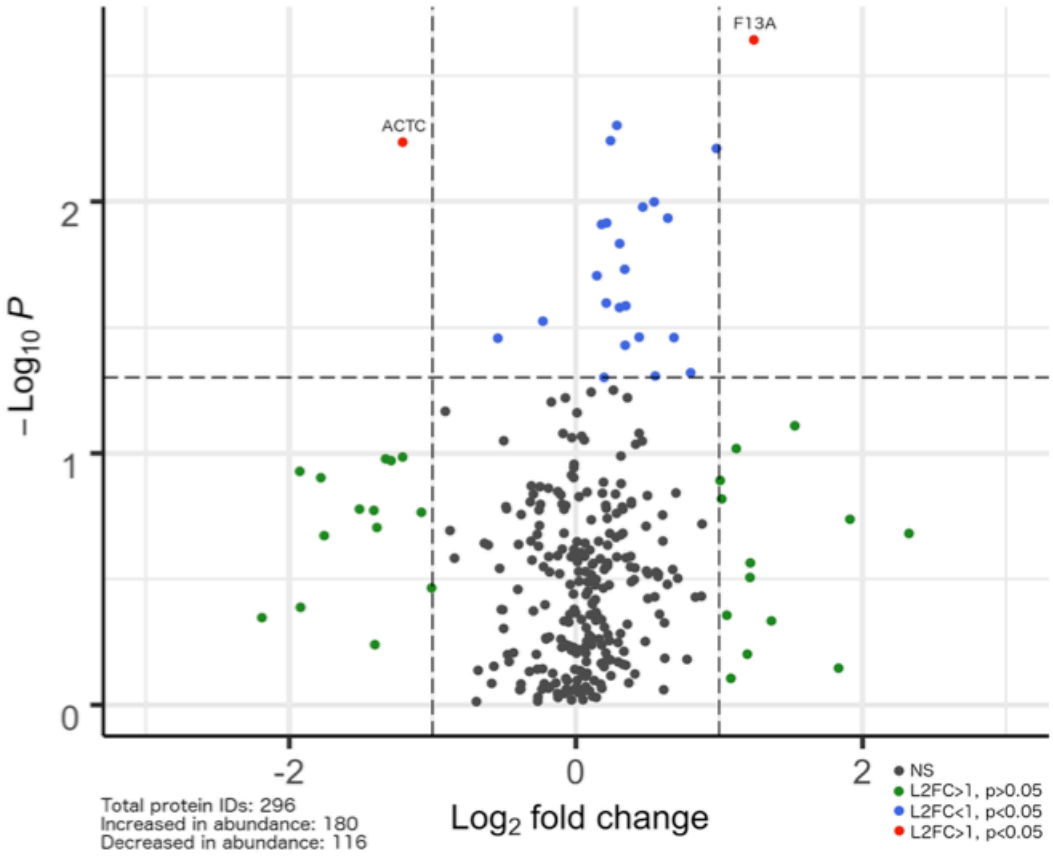


Figure 3.4: Volcano plot using mass spectrometry run 1 data – Lyme disease seropositive vs. all other sera. Proteins on the right of the graph are increased in abundance in Lyme disease seropositive, with proteins decreased shown on the left. A p-value of 0.05 (Log[10]p=1.4) was used as a significance cut-off along with a value of 1 for fold change (Log[2]). Proteins with values above these cut-offs are shown in red.

UniProt ID	Gene	Protein Name	GO - Biological Function	Log^2FoldChange	*p*-value
Proteins increased in abundance in Lyme seropositive sera					
P00488	F13A	Coagulation factor XIII A chain	Blood coagulation	1.241	0.002
P01023	A2MG	Alpha-2-macroglobulin	Blood coagulation, intrinsic pathway	0.981	0.006
P06727	APOA4	Apolipoprotein A-IV	Cellular protein metabolic process	0.801	0.047
P54108	CRIS3	Cysteine-rich secretory protein 3	Defence response	0.685	0.034
P09172	DOPO	Dopamine beta-hydroxylase	Behavioural response to ethanol	0.641	0.012
Q9Y490	TLN1	Talin-1	Cell-cell junction assembly	0.553	0.049
P51884	LUM	Lumican	Cartilage development	0.545	0.01
P04278	SHBG	Sex hormone-binding globulin	Androgen binding	0.467	0.01
Q15848	ADIPO	Adiponectin	Brown fat cell differentiation	0.441	0.034
Q04756	HGFA	Hepatocyte growth factor activator	Proteolysis	0.347	0.025
P23142	FBLN1	Fibulin-1	Blood coagulation, fibrin clot formation	0.343	0.037
P02647	APOA1	Apolipoprotein A-I	Adrenal gland development	0.34	0.018
P05452	TETN	Tetranectin	Bone mineralization	0.304	0.015
P08185	CBG	Corticosteroid-binding globulin	Glucocorticoid metabolic process	0.304	0.026
P02765	FETUA	Alpha-2-HS-glycoprotein	Acute-phase response	0.286	0.004
O75882	ATRN	Attractin	Animal organ morphogenesis	0.241	0.005
P05154	IPSP	Plasma serine protease inhibitor	Blood coagulation	0.214	0.012
P04217	A1BG	Alpha-1B-glycoprotein	Neutrophil degranulation	0.211	0.025
P80108	PHLD	PI-glycan-specific phospholipase D	Cell migration, sprouting angiogenesis	0.197	0.049
Proteins decreased in abundance in Lyme seropositive sera					
UniProt ID	**Gene**	**Protein Name**	**GO - Biological Function**	**Log^2FoldChange**	***p*-value**
P68032	ACTC	Actin, alpha cardiac muscle 1	Actin filament-based movement	-1.208	0.006
P35542	SAA4	Serum amyloid A-4 protein	Acute-phase response	-0.544	0.035
P04004	VTNC	Vitronectin	Cell adhesion	-0.230	0.030

*(adjusted $p<0.05$)

Table 3.4: Table showing the 22 proteins found to be significantly changed in abundance when comparing Lyme disease seropositive and all other sera.

3.3.3.4. Mass spectrometry run 1: Syphilis positive vs. other sera

As fewer differentially abundant proteins were identified between the Lyme disease seropositive and seronegative groups than expected, further analyses of the infectious control groups was conducted to examine whether greater changes could be identified in the serum proteome of patients with other infections. Sera from syphilis positive patients (n=5) were compared to all other sera, a total of 20 proteins were found be at significantly different abundances between groups. Of these, 9 were increased in abundance in syphilis positive while 11 were decreased in abundance. **Figure 3.5** shows a volcano plot of syphilis positive sera compared to all other sera groups. **Table 3.5** lists all proteins significantly increased or decreased in syphilis positive sera. Of those increased in syphilis, 6 proteins identified had a Log[2]Fold Change of >1, with the greatest increase seen for apolipoprotein L1(APO-A1) (p=2.6E-07, Log[2]Fold Change =1.588). In contrast to the comparison between Lyme disease groups, several proteins were found to be highly differentially abundant between groups (Fold Change >4), with high significance. In general, proteins identified as significantly increased in abundance in syphilis positive were associated with cellular metabolic processes. 2 proteins associated with activation of Mitogen-Activated Protein Kinase (MAPK) were found to be increased, suggesting a host response to stimuli during syphilis infection. Of those proteins found to be decreased in abundance in syphilis positive sera, several were associated with blood coagulation.

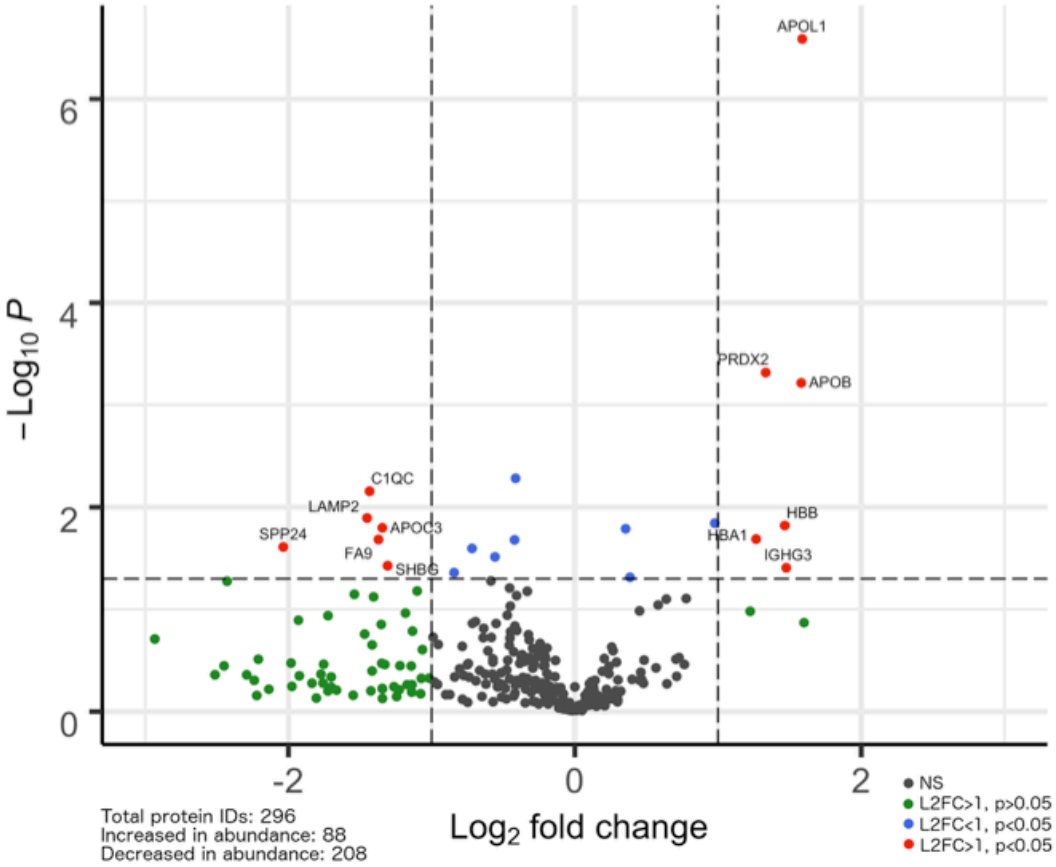


Figure 3.5: Volcano plot using mass spectrometry run 1 data – Syphilis positive vs. all other sera. Proteins on the right of the graph are increased in abundance in syphilis positive, with proteins decreased shown on the left. A p-value of 0.05 (Log[10]p=1.4) was used as a significance cut-off along with a value of 1 for fold change (Log[2]). Proteins with values above these cut-offs are shown in red.

Proteins increased in abundance in Syphilis positive sera					
UniProt ID	**Gene**	**Protein Name**	**GO - Biological Function**	**Log^{2}FoldChange**	***p*-value**
O14791	APOL1	Apolipoprotein L1	Cellular protein metabolic process	1.588	2.6E-07
P00915	APOB	Carbonic anhydrase 1	Biocarbonate transport	1.581	0.001
P01860	IGHG3	IG-heavy constant gamma 3	B cell receptor signalling pathway	1.476	0.039
P68871	HBB	Hemoglobin subunit beta	Bicarbonate transport	1.466	0.015
P32119	PRDX2	Peroxiredoxin-2	Activation of MAPK activity	1.333	4.8E-04
P69905	HBA1	Hemoglobin subunit alpha	Bicarbonate transport	1.266	0.021
P31146	DOPO	Coronin-1A	Actin cytoskeleton organization	0.977	0.014
P01031	APOD	Complement C5	Activation of MAPK activity	0.385	0.049
PP02790	HEMO	Hemopexin	Cellular iron ion homeostasis	0.355	0.016
Proteins decreased in abundance in Syphilis positive sera					
UniProt ID	**Gene**	**Protein Name**	**GO - Biological Function**	**Log^{2}FoldChange**	***p*-value**
Q13103	SPP24	Secreted phosphoprotein 24	Bone remodelling	-2.036	0.024
Q04756	LAMP2	Hepatocyte growth factor activator	Proteolysis	-1.450	0.013
P23142	C1QC	Fibulin-1	Blood coagulation	-1.433	0.007
P00740	FA9	Coagulation factor IX	Blood coagulation	-1.371	0.021
P02656	APOC3	Apolipoprotein C-III	Cholesterol efflux	-1.344	0.016
P04278	SHBG	Sex hormone-binding globulin	Androgen binding	-1.307	0.037
Q15848	ADIPO	Adiponectin	Brown fat cell differentiation	-0.843	0.044
P48740	MASP1	Mannan-binding lectin serine protease 1	Complement activation	-0.718	0.025
P15169	CBPN	Carboxypeptidase N catalytic chain	Bradykinin catabolic process	-0.558	0.031
P02760	AMBP	Protein AMBP	Cell adhesion	-0.421	0.021
P22792	CPN2	Carboxypeptidase N subunit 2	Protein stabilization	-0.414	0.005

*(adjusted $p<0.05$)

Table 3.5: Table showing the 20 proteins found to be significantly changed in abundance when comparing syphilis seropositive and all other sera.

3.3.7. Mass spectrometry run 1: Leptospirosis positive vs. other sera

Of the groups included in mass spectrometry run 1, those from the leptospirosis positive group showed the greatest difference in serum proteome. A total of 146 proteins were found to be in significantly different abundance in the leptospirosis positive group, over 50% of all proteins identified between groups. 121 proteins were found to be significantly increased in abundance in leptospirosis positive but only 25 were significantly decreased. Of those increased, 96 had a Log[2]Fold Change >1, while of those decreased, 12 had a Log[2]Fold change <-1. The mass spectrometry analysis of leptospirosis sera therefore shows a marked change in proteome, likely reflecting various host-responses to leptospiral infection and the systemic, multi-organ nature of the disease. **Figure 3.6** shows a volcano plot of leptospirosis positive sera against all other sera groups. Differentially abundant proteins in leptospirosis positive sera were associated with a wide range of biological functions. **Tables 3.6 and 3.7** show the top 20 proteins significantly increased and decreased in leptospirosis over all other sera respectively. Of those significantly increased were carbonic anhydrase III, a muscle-specific protein that is usually own found at very low concentration in serum. Several acute phase proteins including serum amyloid A2, c-reactive protein and lithostathine-1-alpha were found to be increased in abundance in leptospirosis positive sera. Proteins decreased in leptospirosis sera were found to have a wide range of biological functions. To better understand these changes, together with potential patterns in protein change, the mass spectrometry data from leptospirosis positive sera was compared to Seronegative as a negative control using Ingenuity Pathway Analysis. IPA requires a minimum number of protein/gene inputs that were not met for other proteomic comparisons. For each biological function or pathway, a statistical quantity is computed called the activation z-score. **Table 3.8** shows the top upstream

regulators identified by increased abundance in target molecules in the mass spectrometry dataset. IL6 (Activation z-score = 2.935) and the presence of lipopolysaccharide (Activation z-score = 2.604) were identified as important upstream regulators of pathways induced in leptospirosis positive sera. To illustrate the wide range of protein abundance between leptospirosis positive and other sera, as measured by mass spectrometry, **figure 3.6** shows normalised abundance data for top differentially abundant proteins between these groups.

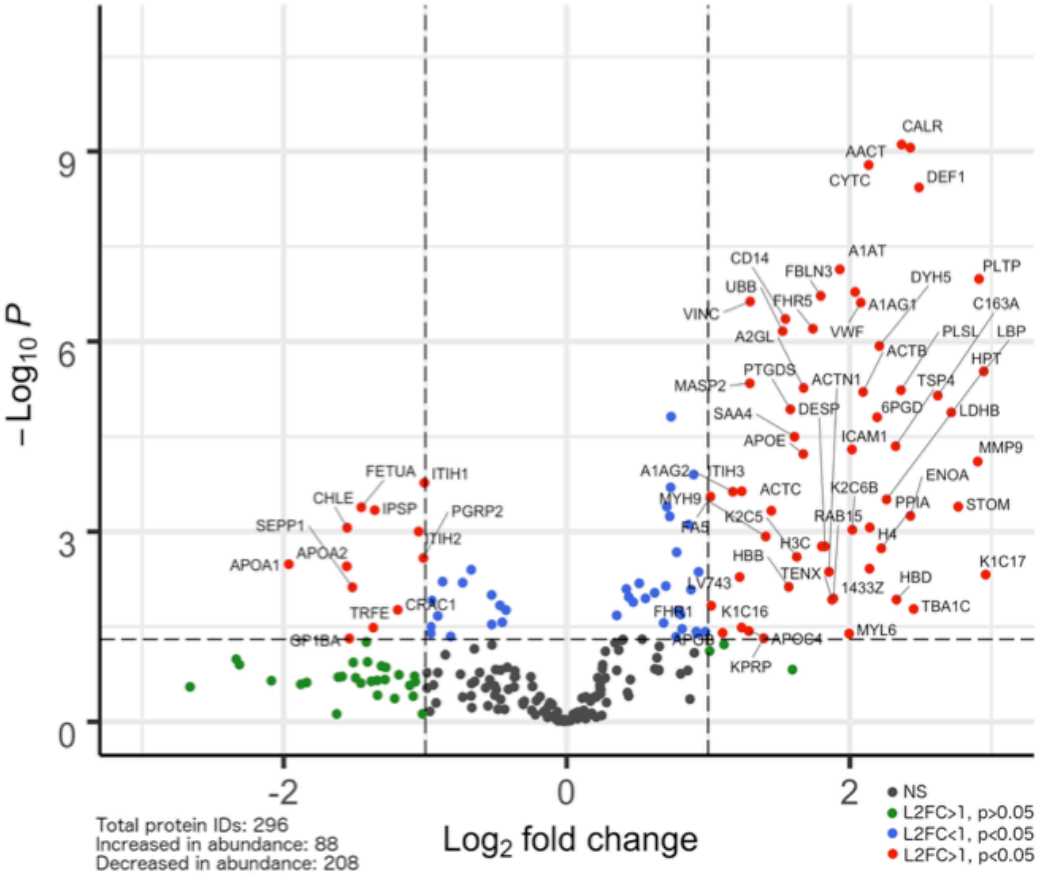


Figure 3.6: Volcano plot using mass spectrometry run 1 data – Leptospirosis positive vs. all other sera. Proteins on the right of the graph are increased in abundance in leptospirosis positive, with proteins decreased shown on the left. A p-value of 0.05 (Log[10]p=1.4) was used as a significance cut-off along with a value of 1 for fold change (Log[2]). Proteins with values above these cut-offs are shown in red.

Proteins increased in abundance in leptospirosis positive sera					
UniProt ID	**Gene**	**Protein Name**	**GO - Biological Function**	**Log²FoldChange**	***p*-value**
P48594	SPB4	Serpin B4	Negative regulation of endopeptidase activity	7.456	2.7E-04
P07451	CAH3	Carbonic anhydrase 3	Response to oxidative stress	7.327	0.006
P06732	KCRM	Creatine kinase M-type	Creatine metabolic process	6.401	4.4E-04
P12883	MYH7	Myosin 7	Adult heart development	6.329	0.007
P0DJI8	SAA1	Serum amyloid A-1 protein	Activation of MAPK activity	5.942	5.0E-05
P05451	REG1A	Lithostathine-1-alpha	Antimicrobial humoral immune response	5.929	6.5E-07
P29034	S10A2	Protein S100-A2	Endothelial cell migration	5.764	0.002
P02144	MYG	Myoglobin	Enucleate erythrocyte differentiation	5.693	0.001
Q9Y279	VSIG4	V-set and IG-domain-containing 4	Complement activation, alt. pathway	5.685	1.0E-08
P0DJI9	SAA2	Serum amyloid A-2	Acute-phase response	4.856	1.1E-06
P02741	CRP	C-reactive protein	Acute-phase response	4.666	2.7E-08
Q5VTE0	EF1A3	Putative elongation factor 1-alpha-like 3	Translation	4.626	2.0E-04
P80188	NGAL	Neutrophil gelatinase-associated lipocalin	Antimicrobial humoral response	4.580	1.4E-06
P08185	CBG	Corticosteroid-binding globulin	Glucocorticoid metabolic process	0.304	0.026
P31947	1433S	14-3-3 protein sigma	DNA damage response	4.556	0.001
P06744	G6PI	Glucose-6-phosphate isomerase	Canonical glycolysis	4.458	2.1E-07
Q9Y6R7	FCGBP	IgGFc-binding protein	Mucosal structure maintenance	4.376	3.6E-08
P22532	SPR2D	Small proline-rich protein 2D	Cornification, epidermis development	4.346	0.005
P06732	KCRM	Creatine kinase M-type	Creatine metabolic process	4.285	4.3E-05
P07355	ANXA2	Annexin A2	Angiogenesis	4.058	0.001

ıdjusted p<0.05)

Table 3.6: Table showing the top 20 proteins found to be significantly increased in abundance in leptospirosis positive when comparing leptospirosis seropositive and all other sera.

Proteins decreased in abundance in Leptospirosis positive sera					
UniProt ID	**Gene**	**Protein Name**	**GO - Biological Function**	**Log^{2}FoldChange**	***p*-value**
P02647	APOA1	Apolipoprotein A-I	Blood vessel endothelial cell migration	-1.963	0.003
P07451	APOA2	Apolipoprotein A-II	Acute inflammatory response	-1.555	0.004
P06276	CHLE	Cholinesterase	Choline metabolic process	-1.552	0.001
P07359	GP1BA	Platelet glycoprotein Ib alpha chain	Blood coagulation	-1.536	0.049
P49908	SEPP1	Selenoprotein P	Brain development	-1.515	0.008
P02765	FETUA	Alpha-2-HS-glycoprotein	Acute-phase response	-1.452	4.1E-04
P02787	TRFE	Serotransferrin-precursor	Cellular iron ion homeostasis	-1.367	0.033
PP05154	IPSP	Plasma serine protease inhibitor	Blood coagulation	-1.357	4.1E-04
Q9NQ79	CRAC1	Cartilage acidic protein 1	Axonal fasciculation	-1.195	0.017
P19823	ITIH2	Inter-alpha-trypsin inhibitor heavy chain	Cellular protein metabolic process	-1.047	0.001
Q96PD5	PGRP2	N-acetylmuramoyl-L-alanine amidase	Antimicrobial humoral response	-1.014	0.003
P19827	ITIH1	Inter-a-trypsin inhibitor heavy chain	Hyaluronan metabolic process	-1.007	1.7E-04
P07996	TSP1	Thrombospondin-1	Activation of MAPK activity	-0.959	0.040
P43652	AFAM	Afamin	Protein stabilization	-0.958	0.032
P02766	TTHY	Transthyretin	Cellular protein metabolic process	-0.951	0.012
P80108	PHLD	PI-glycan-specific lipase D	Cell migration	-0.910	0.021
O75882	ATRN	Attractin	Animal organ morphogenesis	-0.875	0.006
P27169	PON1	Serum paraoxonase	Aromatic compound catabolic process	-0.819	0.045
P48594	SPB4	Serpin B4	Negative reg. of endopeptidase activity	-0.734	0.006
P04180	LCAT	PI-choline-sterol acyltransferase	Cholesterol esterification	-0.674	0.004

*(adjusted p<0.05)

Table 3.7: Table showing the top 20 proteins found to be significantly decreased in abundance in leptospirosis positive when comparing leptospirosis seropositive and all other sera.

Upstream regulator	Molecule type	Activation z-score	Overlap p-value	Target molecules in dataset
IL6	Cytokine	2.935	1E-30	A2M,AGT,ALB,ANXA1,APCS,APOA1,APOB,APOE,C3,CD14,CD163,CFD, CFH,CFP,CLU,COMP,CP,CPB2,CRP,CST3,F12,FGA,FGB,FGG,FN1,GP1BA, GP5,HP,HPX,ICAM1,IGFBP3,IGHM,ITGB3,JCHAIN,KRT14,LBP,LCAT, LCN2,LPA,LRG1,LTF,MMP9,ORM1,PLG,PON1,PPBP,REG1A,S100A7,S100A9, SAA1,SAA2,SAA4,SERPINA1,SERPINA3,SERPINA7,STOM,TF,THBS1,TNC,TT
Lipopoly-saccharide	External	2.604	1.2E-27	ACTC1,ADIPOQ,AGT,ALB,ANXA1,ANXA3,APCS,APOA1,APOA4,APOB,APOC1, APOC2,APOE,AZGP1,C3,C5,CALR,CD14,CD163,CETP,CFB,CFD,CFP,CORO1A,CP, CPB2,CRP,CSF1R,CST3,ENO1,F2,FABP5,FGB,FN1,HP,HSPA8,ICAM1,IGFBP3, IGHG1,IGHM,IGKC,ITIH2,ITIH4,JCHAIN,LBP,LCN2,LDHA,LGALS3BP,LTF,MB, MBL2,MMP9,MYH7,MYH9,ORM1,ORM2,PF4,PKM,PLEK,PLG,PPIA,PROC,PROS1, PTGDS,S100A2,S100A8,S100A9,SAA1,SAA2,SELENOP,SELL,SERPINA1, SERPINA3,SERPINC1,SERPINF1,SERPINF2,SPARC,TF,THBS1,TNC,TPI1, VCL,VSIG4,VTN,VWF
TGFB1	Growth Fac-tor	2.4	1.75E-25	ACTC1,ACTN1,ADIPOQ,AFM,ALB,ALDOA,ANXA2,APOB,APOC2,APOE,C1QA, C1QB,C1QC,C1R,C1S,C2,C3,C4BPA,C4BPB,C5,CAP1,CD14,CD163,CFB,CFD,CFH, CFI,CKM,CLU,COMP,CRP,CSF1R,CST3,DSP,ECM1,ENO1,F12,F13A1,F2,F5,FABP5, FGA,FGB,FGG,FLNA,FN1,GPI,HABP2,ICAM1,IGFBP2,IGFBP3,IGHM,ITGB3, ITIH3,JUP,KNG1,KRT10,KRT14,KRT17,LCAT,LCN2,LDHA,LYVE1,MMP9,MYH7, MYH9,MYL6,PGK1,PKM,PTGDS,SERPINA1,SERPINA3,SERPINF1,SERPINF2, SLC4A1,SPARC,TGFBI,THBS1,TNC,VCL,VWF
HNF1A	Transcription regulator	1.691	8.16E-25	AGT,AHSG,ALB,AMBP,APCS,APOA2,APOB,APOC3,APOH,APOM,C1S,C2, C4BPA,C5,C8A,C8B,C8G,C9,CPB2,CPN1,CRP,F11,FGA,FGB,GC,HABP2,HGFAC, HPX,ITIH4,KNG1,LBP,LCAT,PGK1,PLG,PROC,PZP,SERPINA1,SERPINA10, SERPINA7,SERPING1,TPI1,TTR,VTN
OSM	Cytokine	4.551	2.9E-20	A2M,ACTB,ACTC1,ADIPOQ,AGT,AHSG,ALB,APOA1,APOA2,APOB,APOC1, APOC2,APOC3,APOM,C1S,C4A,C4B,CA1,CLU,CST3,F2,FGA,HP,IGFBP3,ITGB3, ITIH3,KRT16,KRT17,MMP9,MYH7,PGK1,PRDX2,PTGDS,SHBG,TF

Table 3.8: Top upstream regulators identified in leptospirosis positive sera when compared with other sera based on mass spectrometry run 1 data using IPA. Upstream regulators are predicated by increased or decreased abundance of marker proteins. The proteins listed on the right of the table were present in the data set and conformed to the presence of predicted upstream regulated. The z-score is a measure of the conformity to the model based on the dataset.

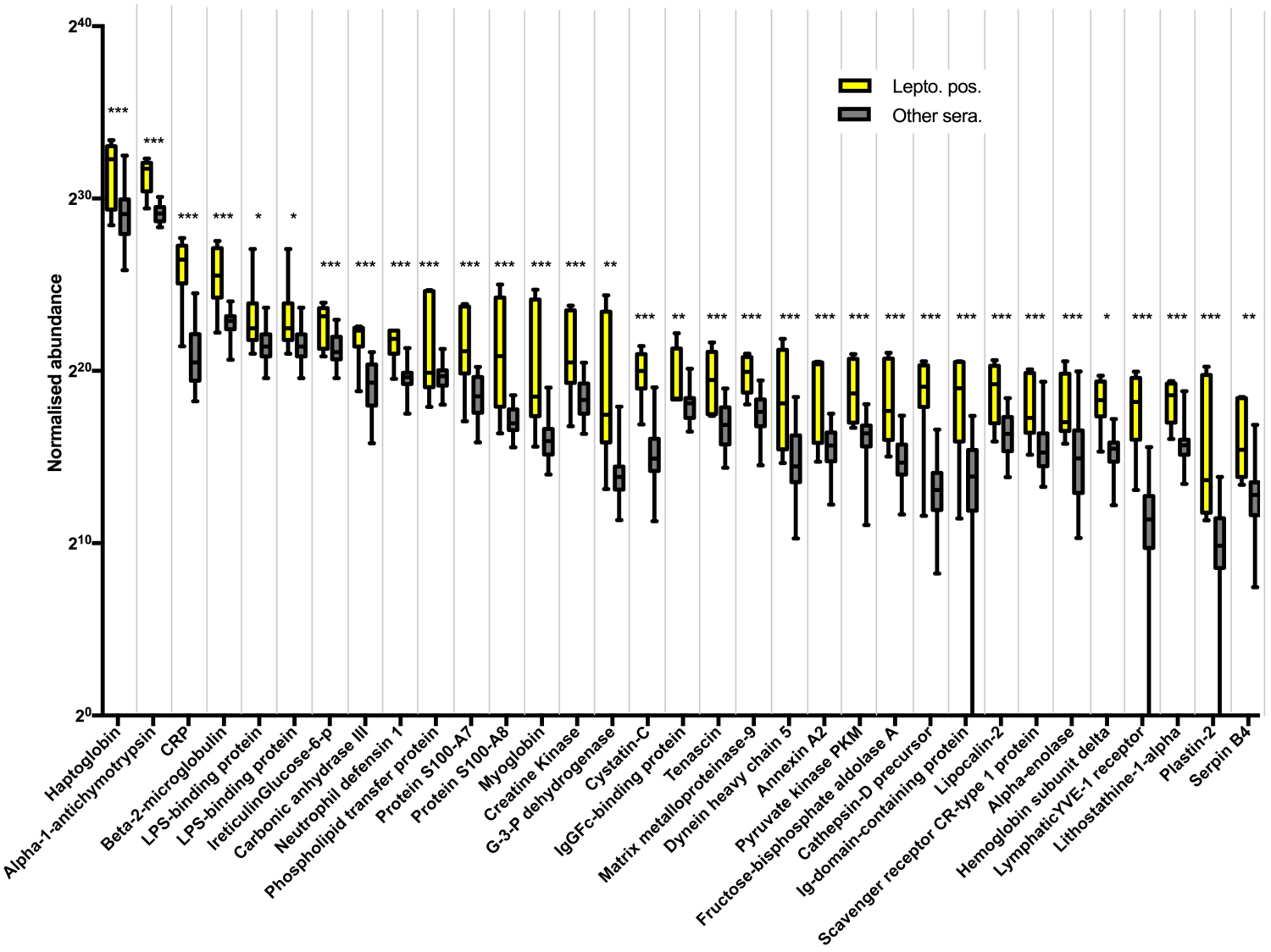


Figure 3.7: Graph showing normalised abundance data for top differentially abundant proteins when comparing leptospirosis positive sera to that of all other sera groups, as measured by quantitative mass spectrometry.

3.3.4. Label free mass spectrometry run 2

3.3.4.1. Overview

As very few proteins were identified as present at different abundances between Lyme disease seropositive and seronegative groups, a second mass spectrometry run was undertaken to further confirm these findings. Due to the source of Lyme disease samples in the initial run (residual sample from diagnostic testing), their storage history and freeze-thaw count were unknown and therefore there was possibility of protein degradation in these samples. In this second mass spectrometry run, samples acquired from the Czech Republic, a country with endemic Lyme disease, were used. The following sample sets were compared: Czech Lyme disease seropositive (n=8), Czech Seronegative (n=8) and normal/healthy donors (serum) (n=8). **Figure 3.8** shows a PCA plot of the three sample groups included in mass spectrometry run 2. While normal healthy donor cells group and can be seen to be distinct, the Lyme disease seropositive and seronegative groups were found to overlap to a high degree. This indicates that at the resolution available by this method, the serum proteome of Lyme disease seropositive and seronegative are highly similar. No proteins were identified between Lyme disease seropositive and seronegative groups that were above the significance ($p<0.05$) or fold change (log [2] >1) cut-offs.

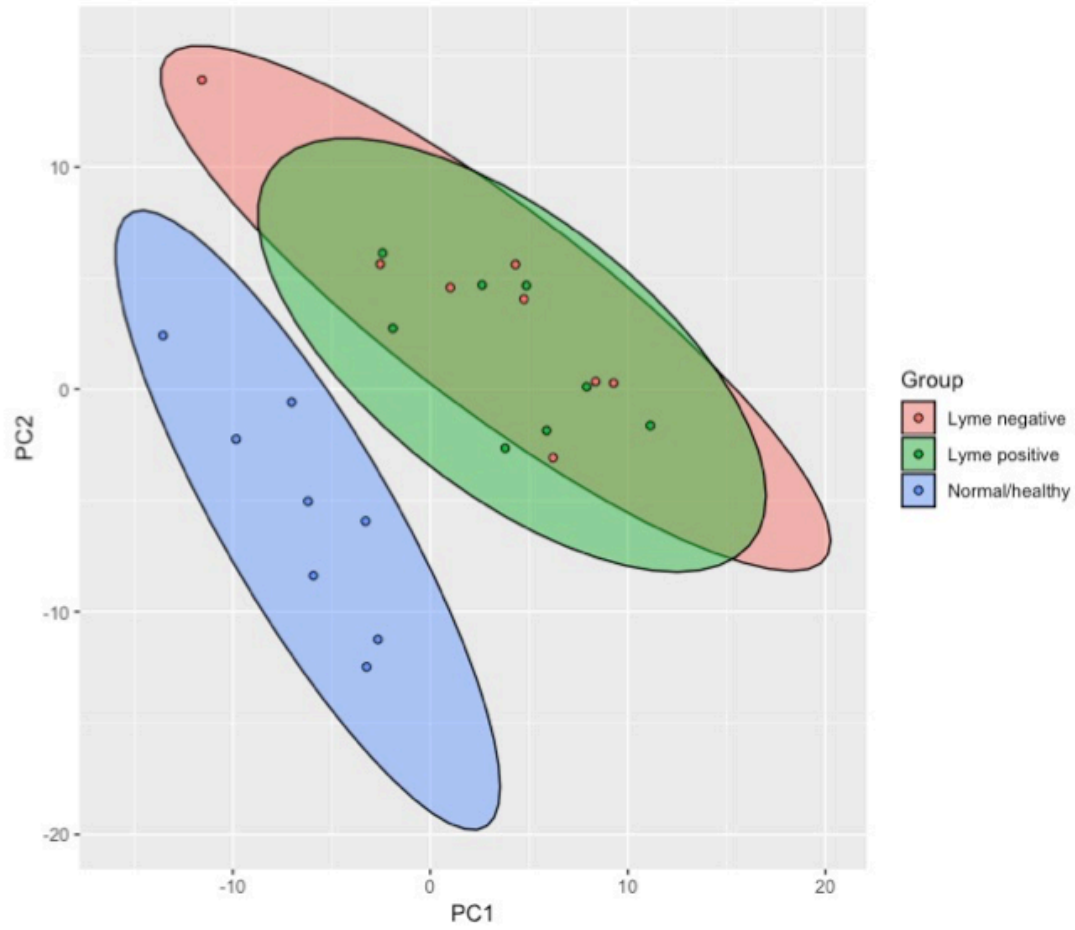


Figure 3.8: Principal Component Analysis (PCA) plot showing normalised data from mass spectrometry run 2. Each point represents a single sample, with colour overlay showing sample groups.

3.3.4.2. Mass spectrometry run 2: Lyme disease groups vs. healthy sera

In contrast to the high similarity of serum proteome seen between Lyme disease seropositive and seronegative, the NHD control group was seen to be consistently distinct at the proteome level from these groups. A total of 72 proteins were found to be at significantly different abundance in the Lyme disease seropositive group when compared directly to the NHD controls. Of these, 41 were increased in abundance in Lyme disease seropositive, with 31 decreased. **Table 3.9** lists the top proteins significantly increased or decreased in Lyme disease seropositive. Of the top 20 proteins increased in Lyme disease seropositive, at least 18 (90%) had a known function related to immunity. Comparisons of Seronegative to NHD controls showed a largely similar pattern.

Proteins increased in abundance in Lyme seropositive sera					
UniProt ID	**Gene**	**Protein Name**	**GO - Biological Function**	**Log^{2}FoldChange**	***p*-value**
P01624	KV315	Immunoglobulin kappa variable 3-15	Complement activation	3.41	0.0003
P04433	KV311	Immunoglobulin kappa variable 3-11	Complement activation	2.35	0.032
O43866	CD5L	CD5 antigen-like	Apoptotic process	2.26	0.009
P01591	IGH	Immunoglobulin J chain	Adaptive immune response	1.98	0.04
P06310	KV230	Immunoglobulin kappa variable 2-30	Brain development	1.96	0.035
P017805	HV307	Immunoglobulin heavy variable 3-7	B cell receptor signaling pathway	1.88	0.007
Q9Y2I7	FYV1	1-phosphatidylinositol 3-phosphate 5-kinase	Intracellular signal transduction	1.78	0.029
P01871	IGHM	Immunoglobulin heavy constant mu	Adaptive immune response	1.78	0.037
P01619	KV320	Immunoglobulin kappa variable 3-20	Antibacterial humoral response	1.55	0.021
Q13790	APOF	Apolipoprotein F	Cholesterol metabolic process	1.32	0.004
P01615	KVD28	Immunoglobulin kappa variable 2D-28	Complement activation	1.28	0.005
P01825	HV4591	Immunoglobulin heavy variable 4-59	B cell receptor signaling	1.26	0.047
Q9H8L6	MMRN2	Multimerin-2	Cell migration	1.09	0.044
P02751	FINC	Fibronectin	Acute-phase response	1.06	0.003
P01861	IGHG4	Immunoglobulin heavy constant gamma 4	B cell signalling	0.98	0.041
P02788	TRFL	Lactotransferrin	Antibacterial humoral response	0.82	0.012
P80748	LV321	Immunoglobulin lambda variable 3-21	Complement activation	0.71	0.031
P01860	IGHG3	Immunoglobulin heavy constant gamma 3	B cell signalling	0.70	0.046
P01772	HV333	Immunoglobulin heavy variable 3-33	NB cell signalling	0.69	0.044

*(adjusted p<0.05)

Table 3.9: Table showing the top 20 proteins found to be significantly increased in abundance in Lyme disease sera when comparing early Lyme to normal healthy donor sera

Proteins decreased in abundance in Lyme seropositive sera					
UniProt ID	**Gene**	**Protein Name**	**GO - Biological Function**	**Log^{2}FoldChange**	***p*-value**
P02671	FIBA	Fibrinogen alpha chain	Adaptive immune response	-3.50	0.003
P07737	PROF1	Profilin-1	Actin cytoskeleton organization	-2.8	0.002
P30447	1A23	HLA class I histocompatibility antigen	Antigen processing	-1.61	0.003
P21333	FLNA	Filamin-A	Actin crosslink formation	-1.60	0.041
Q9UHG3	PCYOX	Prenylcysteine oxidase 1	Prenylated catabolic processy	-1.41	0.008
P02652	APOA2	Apolipoprotein A-II	Acute inflammatory response	-1.29	0.007
P68133	ACTS	Actin, alpha skeletal muscle	Actin crosslink formation	-1.25	0.006
Q15166	PON3	Serum paraoxonase/lactonase 3	Aromatic compound catabolic process	-1.18	0.031
P08709	FA7	Coagulation factor VII	Animal organ regeneration	-1.01	0.009
Q15582	BGH3	TGF-beta-induced protein ig-h3	Angiogenesis	-1.00	0.001
P0C0L5	CO4B	Complement C4	Complement activation	-0.92	0.042
P23528	COF1	Cofilin-1	Actin cytoskeleton organization	-0.82	0.026
Q15848	Q15848	Adiponectin	Brown fat cell differentiation	-0.80	0.043
P01019	ANGT	Angiotensinogen	Activation of phospholipase C activity	-0.69	0.042
P08514	ITA2B	Integrin alpha-IIb	Cell-matrix adhesion	-0.66	0.022
Q13201	MMRN1	Multimerin-1	Blood coagulation	-65	0.041
P22352	GPX3	Glutathione peroxidase 3	Cellular response to oxidative stress	-.059	0.049
P08294	SODE	Extracellular superoxide dismutase	Cellular response to oxidative stress	-0.56	0.041
P16070	CD44	CD44 antigen	Cartilage development	-.050	0.040
P08185	CBG	Corticosteroid-binding globulin	Glucocorticoid metabolic process	-0.049	0.046

*(adjusted $p<0.05$)

Table 3.10: Table showing the top 20 proteins found to be significantly decreased in abundance in Lyme disease sera when comparing early Lyme to normal healthy donor sera

3.4. DISCUSSION

3.4.1. Mass spectrometry run 1: Lyme serum proteome

The aim of this study was to use label-free quantitative mass spectrometry as an exploratory method of identifying novel biomarkers of Lyme disease in human sera. Two mass spectrometry runs were undertaken using serum and plasma samples from patients with Lyme together with related-disease controls and normal healthy donor samples. During the initial discovery run it was found that plasma samples, including the normal healthy donor sample set, were unable to be normalised for comparison to the Lyme and other control groups due to inherent differences introduced during sample preparation. This left the first mass spectrometry effectively with no negative control. Comparisons were made against Seronegative samples as a proxy negative control; however, the nature of these samples meant that the donor had been subjected to routine Lyme disease serological testing and, whilst testing negative, they presumably had another illness with broad symptoms that led to referral for Lyme testing.

When compared directly following mass spectrometry run 1, the Lyme disease seropositive and seronegative groups were found to have highly similar proteomes. A total of 12 proteins were found in significantly different abundances between these groups, with 9 increased in seropositive and 3 decreased. The majority of these changes were relatively small (Log[2]Fold Change between -1 and 1). Lipocalin 2 (LCN2), also known as neutrophil gelatinase-associated lipocalin (NGAL), was found to have the largest difference between these groups (Log[2]Fold Change=2.323, p=0.033), and was taken forward for further analysis in Chapter 4. LCN2 is an innate

immune protein released in particular by neutrophils, epithelial cells and macrophages, and has previously been identified as a biomarker of various processes including inflammation, infection and acute kidney damage(219-221). LCN2 is a 25-kDa secretory glycoprotein, a member of the lipocalin family: a group of proteins that transport small hydrophobic molecules. LCN2 is best described role is it's ability to capture bacterial siderophores, often produced by extracellular bacterial species, ultimately depriving the bacterial cells from obtaining iron. *LCN2*-deficient mice have been shown to be more prone to infection and sepsis (222). Interestingly, the Lyme pathogen *Borrelia burgdorferi* does not require host iron to survive and has been shown to not alter gene expression in the presence of iron chelators. Therefore the impact of an increase in serum lipocalin 2 in host defence from *Borrelia* is difficult to explain without further study. One possibility is that LCN2 is produced in the host as a response to detection of bacterial presence but, due to a lack of bacterial siderophores to bind, the protein builds up in abundance. LCN2 has also previously been identified as increased in skin lesions caused by psoriasis, an autoimmune disease of the skin. Neutralisation of LCN2 led to decrease in inflammation, neutrophilic infiltration and improved epidermal hyperplasia in a murine model (223). LCN2 has also been previously associated with acute kidney injury, caused by leptospirosis infection (224). In a study of host responses in a mouse model to infection by *B. burgdorferi,* LCN2 was identified as being significantly increased (225). In this study, LCN2 was also found to be elevated in leptospirosis positive samples. Given the pathogenesis of Lyme and the associated erythema migrans skin rash, it is possible that LCN2 plays a role in the inflammatory processes involved in Lyme disease. Further study is required to better understand the involvement of LCN2 in Lyme and why it may be increased in Lyme disease patient sera. Regardless of role

or function in Lyme disease, LCN2 was identified as the protein most changed in abundance when comparing Lyme disease seropositive and seronegative and was therefore taken forward as a protein of interest in Chapter 4.

The cross-linked envelope protein, Small proline-rich protein 2D (SPRR-2D), was found to have the second largest abundance difference between Lyme disease seropositive and seronegative by mass spectrometry (Log[2]Fold Change=1.363, p=0.027). At the time of writing, SPRR-2D has not been implicated in Lyme disease Small-proline rich proteins are a group of structural proteins that constitute cornified cell envelope precursors(226). They are therefore involved in the development of epithelial layers, and have been shown to be increased during injury and wound healing. An increased presence of SPRR-2D in Lyme disease seropositive sera may therefore reflect the host skin damage caused by borrelial colonisation at the site of tick-bite. Further structural proteins involved in cell cytoskeletal development, tubulin alpha-4A chain and actin alpha cardiac muscle 1, were found to be lower in abundance in Lyme disease seropositive over seronegative. It is also important to be wary of the presence of skin-associated proteins in mass spectrometry data sets. The presence of the protein keratin, the predominant cytoskeletal protein in epithelia of vertebrates, is often an indication of contamination introduced during sample processing, either during collection or during preparation for mass spectrometry (227). The protein is also highly resistant to digestion by trypsin and pepsin, meaning they are likely to remain at high abundance in the sample. SPRRs and tubulin-subunits are found at much lower abundance and their presence is less likely to be due to simple contamination; however, further work is required understand their presence at differing abundances in host serum.

Fewer differences in protein abundance were detected between Lyme disease seropositive and seronegative samples than expected. Their proteomes were found to be largely similar, with relatively small differences in abundance found. Of the 296 proteins identified across all groups, only 12 (4.1%) were found to be in significantly different abundances between the Lyme disease groups. There were also few significant differences found when comparing the Lyme disease seropositive group to all other sera. As the healthy control group data could not be successfully normalised to allow comparison, analyses lacked a true negative control and were therefore likely to mask general host immune responses to bacteria that may be increased in all groups in mass spectrometry run 1. It was therefore difficult to conclude why so few changes were seen between Lyme disease groups. Firstly, due to the residual diagnostic nature of the samples used, their freeze-thaw cycle count and cold storage history were no known exactly, and the samples could be of poor quality having undergone protein degradation. Lyme disease group samples had been taken in primary care or local phlebotomy departments and are then sent for centralised testing, a process that can take several days and storage conditions are not always fully controlled. This could explain the relatively low protein identification (296) across groups and the presence of proteins that are resistant to proteolytic degradation, notably lipocalin 2 and SPRR-2D. A second possibility for a lack of differences between the 2 Lyme disease groups is a fundamental error in experimental protocol; however, the degree and wide range of changes seen at the proteome level in leptospirosis positive sera indicate that the methodology was capable of detecting proteome changes. Finally, it is possible that the serum proteome level, as measured by label-free quantitative mass spectrometry, of a patient that tests positive for Lyme by Standard Two-Tier testing is largely

similar to those that test seronegative. How these compare to normal, healthy individuals could not be determined in MSI.

3.4.2. Mass spectrometry run 1: Leptospirosis and syphilis positive control groups

In contrast to the Lyme disease seropositive and seronegative groups, the related-disease controls groups (leptospirosis & syphilis positive) showed greater differences in serum proteome. While the main aim of the study was to find distinguishing features between the Lyme disease groups, the control groups showed consistent differences in protein abundances. Some proteins identified in the leptospirosis positive group were taken forward for further analysis in the subsequent chapter. This served as a proof-of-concept to show the ability of mass spectrometry based studies to identify protein biomarkers, but also gave an insight into the host response to *leptospiral* infection.

Syphilis positive sera were found to have 21 proteins (7.1%) at significantly different levels when compared to all other groups. The protein apolipoprotein A1 (APO-A1) was found to be the most increased in abundance when comparing syphilis and other sera groups (Log[2]Fold Change=1.588, p=2.6E-07). The protein acts as the main lytic component in the trypanosome lytic factor (TLF), a component of innate immunity in primates that is characterised by an ability to kill extracellular protozoon pathogens(228). It should be noted that apolipoprotein A1 and apolipoprotein A2 are removed by the affinity columns used during sample preparation. Whilst structurally distinct and not targeted by affinity depletion, APO-A1 forms part of the apolipoprotein group of proteins that share similar lipid binding properties. The presence of APO-A1 in significantly differing amounts may be an artefact as a result

of differences in procedure when processing whole blood to serum. This may also account partially for differences in abundances of several blood coagulation factors, including fibulin-1 (C1QC) and coagulation factor IX (FA9) in the mass spectrometry data. Whilst normalisation of samples can better represent the serum proteome, it will not be able to remove proteins present as a result of differing sample preparation techniques. To better understand results in the context of sample preparation in future mass spectrometry studies, all sample groups could be processed from frozen, stabilised whole blood by exactly the same method to ensure abundance differences identified were due to differences in serum proteome.

Of all sample groups compared during mass spectrometry run 1, the leptospirosis positive set showed the greatest differences in abundance of proteins when compared to all other sera. A total of 146 proteins (75%) were found at significantly different abundance from all other sera. Of these, 121 were increased in leptospirosis positive sera with 25 decreased. The proteins represented a wide-range of functions, with many associated with innate and adaptive host immune response. The scale of the change in proteome seen is likely to be a reflection of the systemic, multi-organ nature of leptospiral infection and the associated robust host immune and inflammatory response that contribute to several disease manifestations. The protein serpin B4 was found to be the most increased in abundance when comparing leptospirosis positive against all other sera, and the greatest difference in abundance across the entire sample set (Log[2]Fold Change=7.456, p=2.4E-04). Serpins are a superfamily of proteins with similar structures. Serpin B4 is a potent inhibitor of serine protease, enzymes that cleave peptide bonds in proteins (229). The fact that a protease inhibitor that is itself resistant to degradation is found at high abundance in leptospirosis

positive sera is highly likely to be as a result of overall sample degradation, rather than a difference that would be observable in the host. Carbonic anhydrase III (CAH3) protein was found to be have the second largest fold change between leptospirosis positive and other sera (Log[2]Fold Change=7.327, p=0.006). CAH3 is an enzyme that catalyses the reversible hydration of carbon dioxide. Importantly, the expression of CAH3 is highly tissue-specific, with high levels seen in skeletal and cardiac muscle(230). It is usually found in low abundance in human blood. Due to the large difference in abundance across sample sets, CAH3 was taken forward for further validation in the subsequent chapter. A possible hypothesis for increased abundance in leptospirosis positive sera is that muscle damage, a common symptom of leptospirosis, causes the release of CAH3 into the bloodstream.

Other proteins identified as differentially abundance in leptospirosis positive sera included apolipoprotein A-1 (APO-A1), transthyretin (TTR) and transferrin (TF) during this study were also previously found to be increased in abundance in a study by Tan *et. al (231)* that compared mild to severe leptospirosis sera. APO-A1 (Log[2]Fold Change=-1.963, p=0.003) and APO2 (Log[2]Fold Change=-1.555, p=0.004) represent the largest decrease in abundance in protein when comparing leptospirosis positive to all other sera. Transthyretin was also found to be highly decreased in abundance in leptospirosis (Log[2]Fold Change=-0.951, p=0.012) as was the transferrin precursor (TRFE) (Log[2]Fold Change=-1.367, p=0.033). In total, the Tan study found 5 proteins that were found at significantly different abundance between healthy and mild leptospirosis (231). The presence of all of these proteins in the mass spectrometry run 1 data set, with significant differences in abundances in the

same direction, showed that mass spectrometry methods are capable of reproducibility across studies.

Ingenuity pathway analysis, based on the data from mass spectrometry 1, identified several key upstream regulators in the leptospirosis positive group. The analysis is based on the presence of specific proteins in the dataset at differing abundance across groups. Interleukin 6 (IL-6) was identified by IPA an important upstream regulator in leptospirosis (Activation z-score = 2.935, p=1E-30). IL-6 has been shown to be an important mediator of fever and of the acute phase response in a number of infections (232). IL-6 has been shown to differentiate severe cases of leptospirosis from mild disease, along with IL-10(233), suggesting it plays an important role in the development of the cytokine storm that leads to severe disease (234).

3.4.3. Mass spectrometry run 2: Overview

During mass spectrometry run 1, plasma samples were excluded from the analysis due to the inability of normalisation methods to reliably compare them to serum-based groups. As so few changes were seen between the Lyme disease seropositive and seronegative groups, and 2 of the differentially expressed proteins were highly resistant to proteases, the quality of the residual-diagnostic samples used in the first run was questionable. To address these issues, mass spectrometry run 2 was undertaken.

Sera taken and stored at time of diagnostic testing for Lyme was used in this run, rather than older residual diagnostic samples. It was hypothesised that the samples would be of better quality than those used in mass spectrometry run 1 and that more proteins would be identifiable. A negative control group was also included,

comprising serum samples from healthy normal donors, recruited locally at PHE Porton. Despite the assumed higher sample quality, fewer proteins (n=285) were successfully quantified across the three sample groups included in mass spectrometry run 2 than in the first run. Of these, 205 were detected in both mass spectrometry run 1 and 2. 91 proteins were only identified in mass spectrometry run 1 only. 80 proteins were identified in mass spectrometry run 2 only. Data from mass spectrometry run 2 showed that sample quality was unlikely to be a limiting factor during the first run, and that a lack of changes in protein abundance observed between Lyme disease seropositive and seronegative is more likely to be due to high similarity in serum proteome between these groups.

3.4.4. Mass spectrometry run 2: Lyme disease positive vs. NHD

In a comparison of Lyme disease positive against normal/healthy donor samples following mass spectrometry run 2, 72 proteins were found at significantly different abundances between groups. Of 41 proteins found to be significantly increased, the majority were directly associated with immune processes. Of the top 20 increased proteins, 13 comprised immunoglobulin subunits, chains or precursor proteins. Other increased proteins in the Lyme disease seropositive group included complement system subunits and activators. The protein found to be at the lowest abundance in relation to the NHD controls was fibrinogen alpha chain. As fibrinogen should largely be removed by affinity columns in the sample preparation process, the presence of this fibrinogen chain again points towards potential influence of sample preparation on results. Several proteins associated with actin cytoskeleton organisation were identified including profilin-1 and filamin-A. This concurs with the results from MSI where several mediators of actin rearrangement were also identified as being

disrupted. The data suggests that host actin rearrangement may be influenced by borrelial infection; however, based on these datasets any consistent changes or mode-of-action are difficult to interpret without further study of the host response.

In general, comparison of Lyme disease seropositive to NHD controls shows a serum proteome that may reflect general host response to bacterial infection. An increase in several immune mediators is seen, as well as increased abundance of immunoglobulin subunits. Without infectious-disease controls present in this run, it was not possible to identify markers that may be more specific to early Lyme disease. As STT testing in the Czech Republic is part of a referral system, it is likely that those tested for Lyme disease have some sort of illness, even if they test seronegative for Lyme disease

3.4.5. Conclusions

The sensitivity of STT in early Lyme has been discussed extensively before (181, 235, 236). Most studies concluded that in early disease, when the human host is still developing an immune response against the pathogen, the level of *Borrelia*-specific antibodies in serum are too low to be readily detectable. The data from the two mass spectrometry runs undertaken in this chapter show that the serum proteome in early Lyme disease may also not be significantly altered. This may reflect the pathogenesis of the disease. In early infection, the pathogen is localised to the skin and circulating levels of bacteria are negligible. It is therefore probable that changes to the serum proteome at this stage would be relatively small.

CHAPTER 4: ANALYSIS OF BIOMARKERS FOR SPIROCHAETAL INFECTION IN HUMAN SERA

4.1. INTRODUCTION

In chapter 3, two separate label-free mass spectrometry experiments were undertaken on Lyme disease patient serum samples, together with infectious disease controls and normal healthy donors. A comparison of Lyme disease seropositive (as diagnosed by Standard Two-Tier testing) sera to Seronegative revealed a small number of differentially abundant proteins between groups. Considerations of initial mass-spectrometry experimental design and extensive follow-up analyses are required before assuming the value of any potential biomarker identified. Biomarkers identified as differentially abundant by mass spectrometry do also not necessarily directly translate to reliable measurement by more routine methods including by Western blot and Enzyme-Linked Immunosorbent Assay (ELISA). Both of these methods utilise antibodies against ligands (in this case proteins) of interest in order to indirectly measure abundance. Discrepancies in mass spectrometry and methods such as ELISA can exist for several reasons including bias introduced during sample preparation for mass spectrometry (237), presence of host-antibodies in biological samples interfering with ELISA binding and any issues surrounding specificity of the primary and secondary antibodies used in such tests. Therefore optimisation and validation of ELISA and WB protocols for use in sera is essential.

In this study, proteins previously identified in chapter 3 as differentially abundant in Lyme disease seropositive sera over Seronegative by mass spectrometry were further analysed in patient sera by WB and ELISA. The aim of this study to was to further

investigate earlier mass spectrometry by common laboratory methods and to assess the ability of WB and ELISA techniques to accurately quantify the abundance of proteins of interest in human serum. To further understand the association between mass spectrometry data and subsequent serological analysis, proteins of interest identified in the leptospirosis positive controls were also included for study.

4.1.1. Hypothesis

The main hypothesis of this chapter was that disease-specific protein abundance changes identified in serum by label-free quantitative mass spectrometry could be detected and quantified by common laboratory techniques including WB and ELISA.

4.2. MATERIALS AND METHODS

Full materials and method details are provided in Chapter 2. **Table 4.1** gives a summary of all sample groups included in this study for analysis by WB and ELISA methods. All available clinical information and diagnostic testing results are described for each sample in Chapter 2 (2.1.1).

Sample set	Number	Criteria for Inclusion	Sample type	Provider
UK Lyme positive (IDs: LP1-LP26)	26	Sera submitted for Lyme disease testing that are seropositive and have a clinical history / symptoms consistent with early Lyme disease.	Serum - Retrospective; residual diagnostic samples	RIPL, PHE Porton
UK Lyme negative (IDs: LN1-LN17)	17	Sera submitted for Lyme disease testing that are seronegative and have a clinical history that is not consistent with early Lyme disease.	Serum - Retrospective; residual diagnostic samples	RIPL, PHE Porton
Czech Lyme positive (IDs: CLP1-CLP26)	26	Sera submitted for Lyme disease testing that are seropositive and have a clinical history / symptoms consistent with early Lyme disease.	Serum - Retrospective; residual diagnostic samples	Hospital Ceske Budejovice, Czech Republic
Czech Lyme negative (IDs: CLN1-CLN24)	24	Sera submitted for Lyme disease testing that are seronegative and have a clinical history that is not consistent with early Lyme disease.	Serum - Retrospective; residual diagnostic samples	Hospital Ceske Budejovice, Czech Republic
Syphilis positive (IDs: SYP1-SYP10)	10	Positive diagnostic samples consistent with recent/active syphilis (seronegative for Lyme disease by STT testing)	Serum - Retrospective; residual sera provided by NHSBT	BRD, PHE Colindale
Leptospirosis positive (IDs: LEP1-LEP18)	18	Positive diagnostic samples consistent with acute leptospirosis (seronegative for Lyme disease by STT testing)	Serum - Retrospective; residual sera provided by NHSBT	RIPL, PHE Porton. BRD, PHE Colindale
Normal Healthy Donors (IDs: H1-H30)	30	Normal healthy donor sera (tested seronegative for Lyme disease by routine testing protocol) from deploying military personnel	Serum - NHD serum aliquots	RIPL, PHE Porton

Abbreviations: RIPL Rare and Imported Pathogens Laboratory, PHE Porton; BRD Bacteriology Reference Department, PHE Colindale; NHSBT National Health Service Blood and Transfusion; UoL - University of Liverpool.

Table 4.1: Sample groups included in WB and ELISA analyses.

4.3. RESULTS

4.3.1. Lyme disease

During mass spectrometry, 7 proteins were identified as differentially abundant with a log[2]fold change greater than 1 in Lyme disease seropositive sera as compared to Seronegative sera. Very low abundance and/or blood-coagulation related proteins were discounted and the proteins lipocalin 2 (LCN2), tubulin alpha (TUBA4A) and small proline rich protein 2D (SPRR2D) were taken forward for further analysis by WB and/or ELISA.

4.3.1.1. Lipocalin-2 Western blot

Mass spectrometry identified the protein lipocalin 2 (LCN2) as significantly increased in abundance in Lyme disease seropositive sera when compared to Seronegative (Log^2 Fold Change =2.323, p=0.33). **Figure 4.1** shows the normalised abundance data for LCN2 as measured by mass spectrometry.

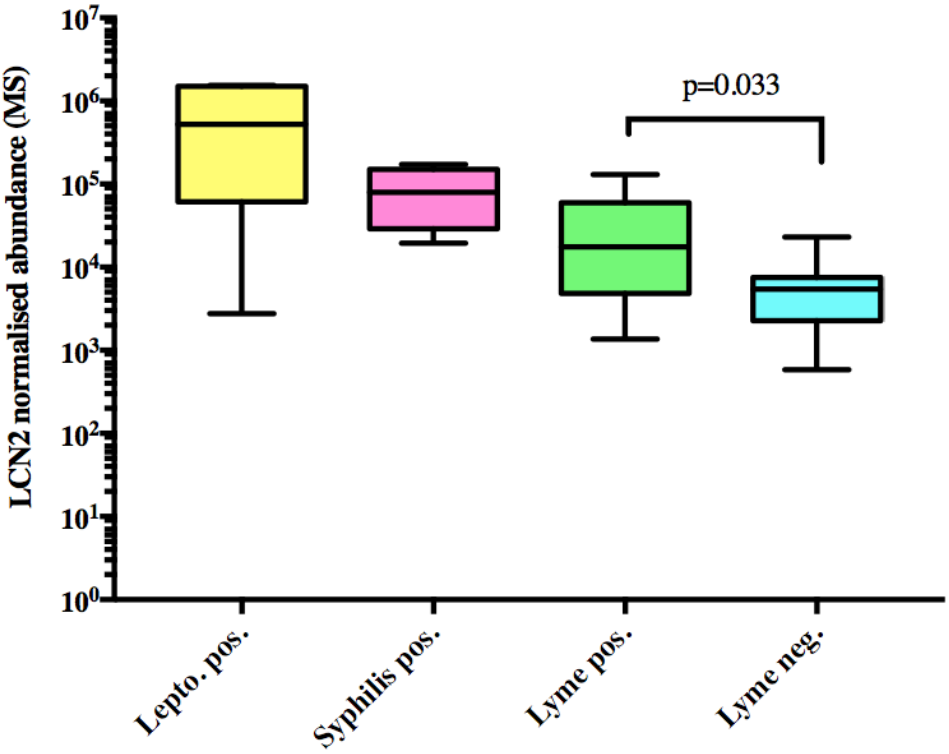


Figure 4.1: Chart showing normalised protein abundance data for lipocalin 2 in human sera by label-free quantitative mass spectrometry. Error bars represent standard error of the mean (SEM). Sample groups; Leptospirosis positive n=5, syphilis positive n=5, Lyme disease seropositive (by STT) n=15, Seronegative (by STT) n=15.

In order to corroborate these findings, a Western blot method for LCN2 was performed on additional serum samples not used in the mass spectrometry discovery run. Serum samples were diluted to 1:50 concentration in BSA 10% buffer, ran using polyacrylamide gel electrophoresis (PAGE) and blotted using a semi-dry Western blot method (Materials and Methods: 3.2.3), using a primary antibody against LCN2 (abcam ab63929). All blots were also stained by SYPRO ruby to allow normalisation by total protein, as described by Eaton *et al* (238) and summarised in Materials and Methods: 3.2.3. **Figure 4.2(A)** shows an example WB image for all sample groups. The expected band size for LCN2 is 22kDA. As predicted, leptospirosis positive and syphilis positive sera showed the highest abundance of LCN2, showing oversaturated bands at the predicted molecular weight. A secondary, unidentified band was observed in several samples. In order to measure differences in LCN2 abundance for Lyme disease seropositive and seronegative groups, 12 samples from each of these groups (n=24) were blotted for LCN2. Each blot contained a negative control (diluent only) and a gel reference sample (LEP-8) to allow comparison between blots. Density of the LCN2 band was measured for each sample using Bio-Rad Image Lab software and results normalised using total protein staining by SYPRO ruby (Materials and Methods: 3.2.4). Results are given as adjusted-relative band density (as compared to a gel reference value = 1). **Table 4.2** shows the results for each sample. The mean relative band density for Lyme disease seropositive sera was 1.059(±0.0363) and for seronegative was 0.927(±0.033). An unpaired *t*-test on the data gave a *p*-value of 0.0409, indicating that a small but significant increase in LCN2 abundance can be seen in the Lyme disease seropositive group.

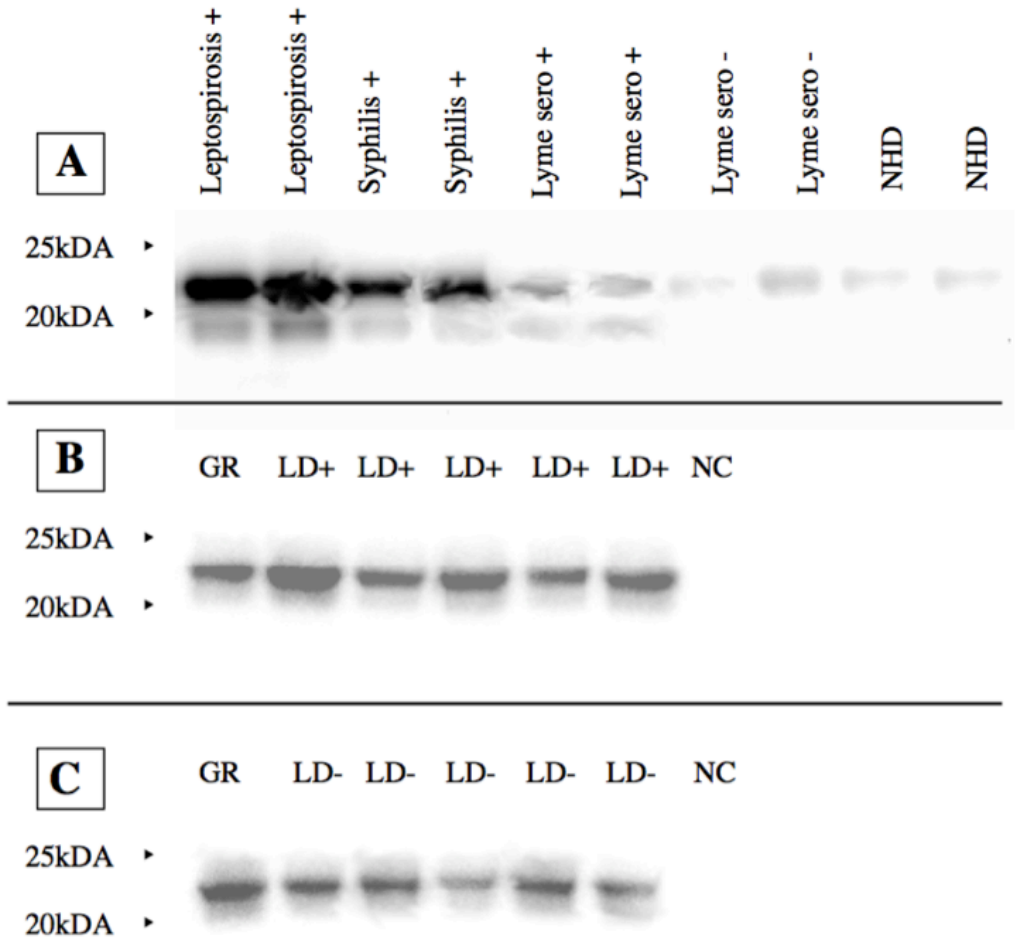


Figure 4.2: Western blots for the protein LCN2 in human sera. A: Blot showing predicted LCN2 bands for leptospirosis positive, syphilis positive, Lyme disease seropositive (by STT), Seronegative (by STT) and normal healthy donor (NHD) sera. B: Blot showing predicted LCN2 band in 5 Lyme disease seropositive serum samples. C: Blot showing predicted LCN2 band in 5 Seronegative serum samples. Primary antibody: Abcam Anti-Lipocalin 2 (ab63929). NC = Negative control (loading buffer only), GR = Gel reference, a highly diluted leptospirosis positive serum sample used on each blot to allow comparison between blots)

Sample ID	Relative band density	Sample ID	Relative band density
C L.Pos. 11	1.324	**C L. Neg. 1**	0.877
C L.Pos. 12	0.983	**C L. Neg. 2**	0.964
C L.Pos. 13	1.013	**C L. Neg. 3**	0.635
C L.Pos. 14	1.164	**C L. Neg. 4**	0.893
C L.Pos. 15	0.943	**C L. Neg. 5**	1.022
C L.Pos. 16	1.033	**C L. Neg. 6**	0.975
C L.Pos. 17	0.983	**C L. Neg. 7**	0.894
C L.Pos. 18	1.235	**C L. Neg. 8**	1.104
C L.Pos. 19	1.022	**C L. Neg. 9**	0.922
C L.Pos. 20	0.892	**C L. Neg. 10**	1.012
C L.Pos. 21	1.01	**C L. Neg. 11**	0.897
C L.Pos. 22	1.103	**C L. Neg. 12**	0.924
Mean	1.059(±0.0363)	**Mean**	0.927 (±0.033)
Unpaired *t*-test	p=0.0409		

Table 4.2: Relative band density for LCN2 in Lyme disease seropositive and seronegative sera (by STT). Lyme disease positive samples are shown on the left hand side of the table and seronegative on the right. All density values are relative to the gel reference band (LEP-8) that was given the value of 1.0. Mean density for each group is shown with 95% confidence interval.

4.3.1.2. Lipocalin 2 ELISA

In order to further confirm these findings by a more sensitive and quantitative method, and to allow the processing of a higher number of samples, an ELISA kit for LCN2 (R&D Human-Lipocalin DuoSet ELISA kit: DY1757) was used. In general, kit protocol was followed as described; however, to optimise the ELISA for use with serum samples, a diluent containing PBS and 10% bovine serum albumin was used (Materials and Methods: 3.3.2). **Figure 4.3** shows a graph of LCN2 serum concentration for all groups, as measured by ELISA. **Figure 4.3(A)** shows all sample

groups separately, with each point representing one serum sample. The mean concentration for each group were as follows: UK Lyme disease positive = 102.9(±30.5) ng/mL (n=26), UK Lyme disease seronegative = 57.37(±21.88) ng/mL (n=17), Czech Lyme disease positive = 92.88(±28.32)ng/mL (n=25), Czech Lyme disease seronegative = 64.5(±17.51)ng/mL (n=24), leptospirosis positive = 418.7(±167.8)ng/mL (n=18), syphilis positive (159.3±51.3) n=10, NHD = 39.64(±11.21)ng/mL (n=30). These data corroborate the previous results of the mass spectrometry analysis, with leptospirosis positive sera showing the highest abundance. **Figure 4.3(B)** shows the combined results for the Lyme disease seropositive groups (UK and Czech Lyme disease seropositive) and for the Seronegative groups (UK and Czech Seronegative). The mean concentration in Lyme disease seropositive was 99.93(±10.04)ng/mL (n=50) and in Seronegative was 63.08(±6.463)ng/mL. Therefore a difference of means of 36.85(±12.64)ng/mL was found between groups with a *p*-value of 0.0035, demonstrating a significant difference in abundance of LCN2 between Lyme disease seropositive and seronegative groups, as measured by ELISA.

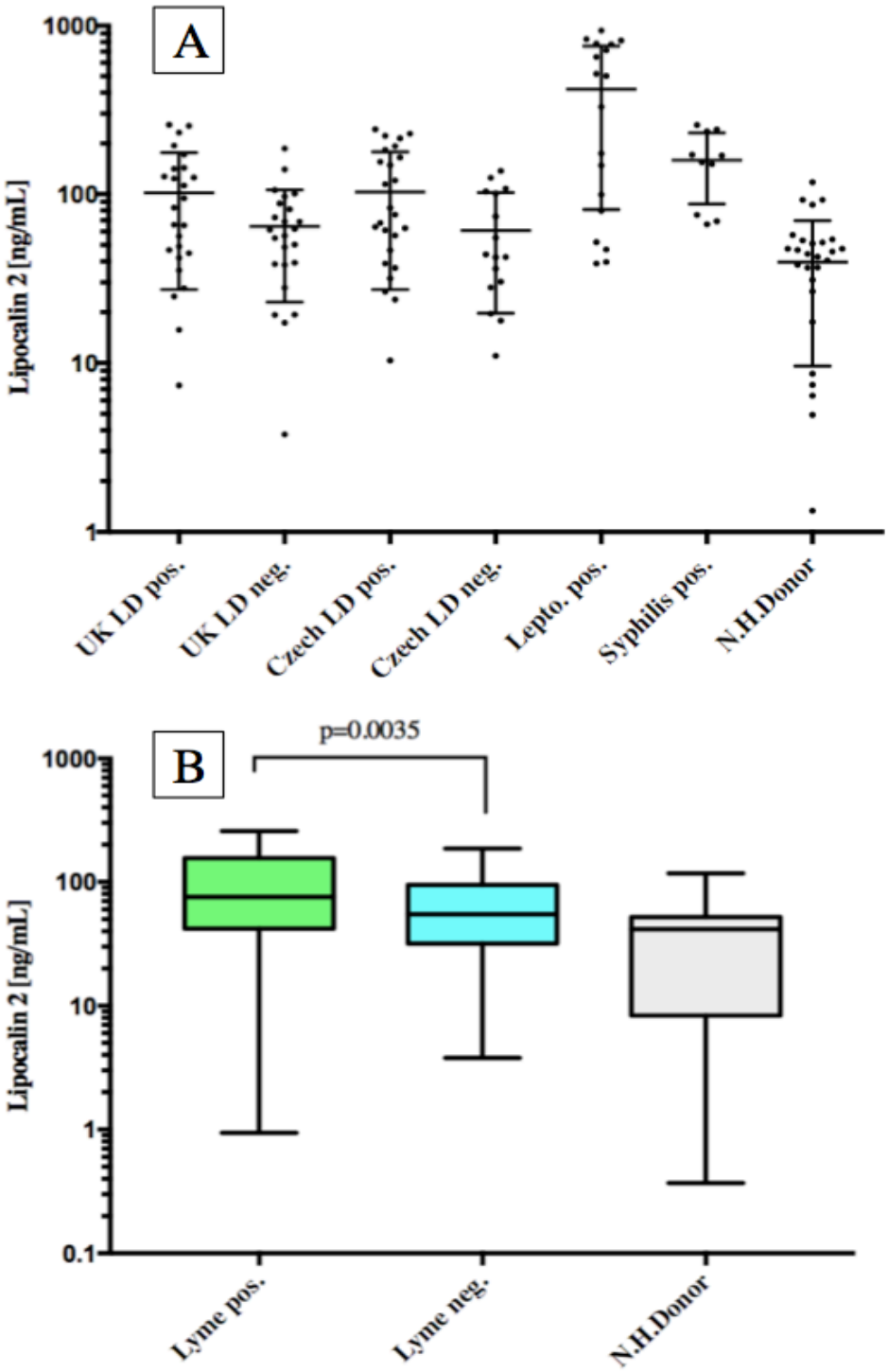


Figure 4.3: Graphs showing Lipocalin 2 abundance as measured by ELISA (R&D Human-Lipocalin DuoSet ELISA kit: DY1757). 5.3(A) shows abundance data for all groups. 5.3(B) shows abundance data for combined Lyme disease positive groups (UK and Czech Lyme disease seropositive) and combined Lyme disease seronegative groups (UK and Czech Seronegative) along with the NHD group. Combined Lyme disease positive n=52, combined Lyme disease seronegative n=50, NH donor n=30.

4.3.1.3. Tubulin-alpha 4A Western blot

Previously, mass spectrometry identified the protein tubulin-alpha 4A (TUBA4A as significantly decreased in abundance in Lyme disease seropositive sera when compared to Seronegative (Log^2 Fold Change =-1.926, p=0.3). **Figure 4.4** shows a graph of normalised abundance for TUBA4A as measured by quantitative mass spectrometry.

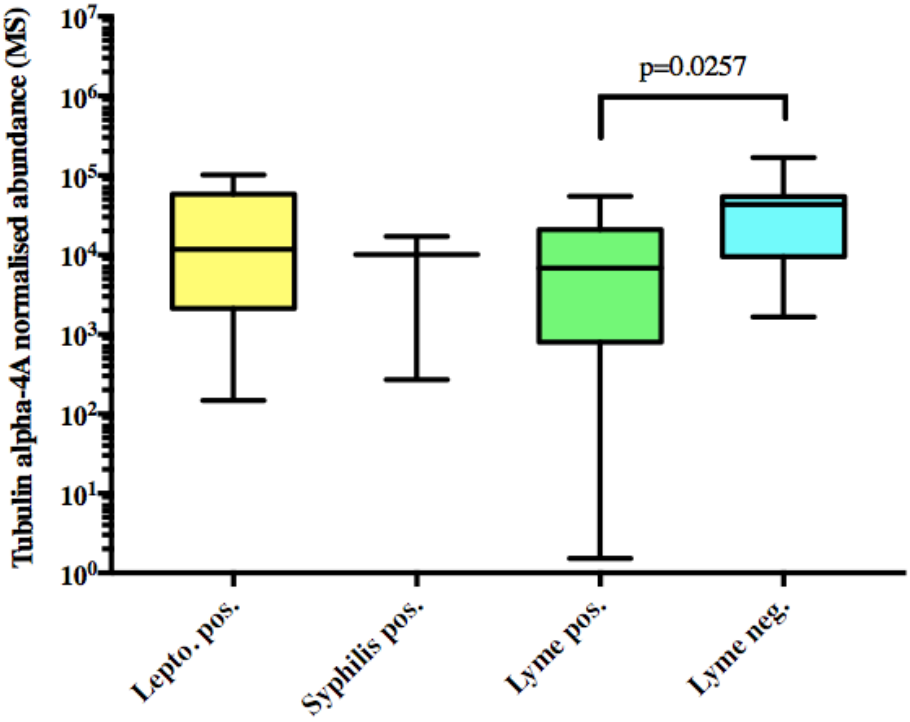


Figure 4.4: Chart showing normalised protein abundance data for tubulin-alpha 4A in human sera by label-free quantitative mass spectrometry. Error bars represent standard error of the mean (SEM). Sample groups; Leptospirosis positive n=5, syphilis positive n=5, Lyme disease seropositive (by STT) n=15, Seronegative (by STT) n=15.

To assess if differences in serum abundance of TUBA4A could be measured and corroborated by WB, serum samples were ran using primary antibodies for tubulin-alpha 4A protein (abcam ab15246). The Western blot procedure as described for LCN2 in section 5.3.1.2 was followed. **Figure 4.5** shows a Western blot for TUBA4A with a range of sample types. By eye, the abundance of the presumed tubulin alpha band (tubulin alpha = 50kda) is similar across sample types. In order to measure protein abundance difference previously identified by mass spectrometry between Lyme disease seropositive and seronegative groups, 12 Lyme disease seropositive and 12 seronegative serum samples not previously used for mass spectrometry were run on gels with a leptospirosis positive gel reference. Samples were normalised using total protein staining by SYPRO ruby and band density was measured against the gel reference (Materials and Methods: 3.2.4). **Table 4.3** shows relative band density for each sample as compared to the gel reference. Only very small differences in band density were observed between sample, and no significant difference in band density was found between groups ($p=0.152$). For all sample groups, extensive smearing of bands and the presence of unexpected bands were observed. Problems were still experienced when using an alternative primary antibody. Presence of bands of lower molecular weight than 50kDA could indicate target protein degradation.

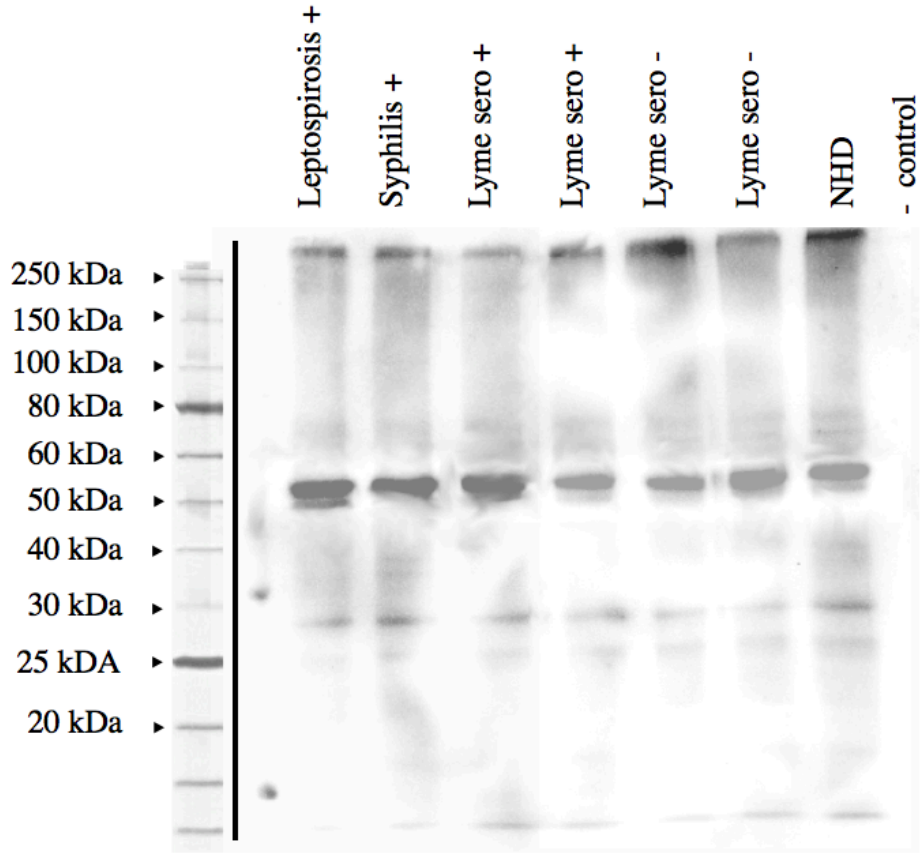


Figure 4.5: Blot showing predicted tubulin-alpha 4A bands for leptospirosis positive, syphilis positive, Lyme disease seropositive (by STT), Seronegative (by STT) and normal healthy donor (NHD) sera. Primary antibody: Abcam Anti-tubulin 4A

Sample ID	Relative band density	Sample ID	Relative band density
C L.Pos. 11	1.012	**C L. Neg. 1**	0.982
C L.Pos. 12	0.953	**C L. Neg. 2**	1.102
C L.Pos. 13	0.981	**C L. Neg. 3**	1.011
C L.Pos. 14	1.114	**C L. Neg. 4**	0.987
C L.Pos. 15	0.843	**C L. Neg. 5**	1.102
C L.Pos. 16	1.033	**C L. Neg. 6**	1.224
C L.Pos. 17	0.991	**C L. Neg. 7**	0.981
C L.Pos. 18	1.011	**C L. Neg. 8**	0.941
C L.Pos. 19	0.989	**C L. Neg. 9**	1.230
C L.Pos. 20	0.897	**C L. Neg. 10**	1.023
C L.Pos. 21	1.001	**C L. Neg. 11**	1.102
C L.Pos. 22	0.809	**C L. Neg. 12**	0.987
Mean	0.969(±0.0242)	**Mean**	1.032(±0.0288)
Unpaired *t*-test	p=0.152		

Table 4.3: Relative band density for TUB4A in Lyme disease seropositive and seronegative sera (by STT). Lyme disease positive samples are shown on the left hand side of the table and seronegative on the right. All density values are relative to the gel reference band (LEP-8) that was given the value of 1.0. Mean density for each group is shown with 95% confidence interval.

4.3.1.4. SPRR-2D, F13A and other low abundance proteins

Several proteins found to be in significantly different abundances between Lyme disease seropositive and Seronegative by mass spectrometry were identified at relatively low abundance. To investigate if these changes were detectable by Western blot, additional serum samples (not used in mass spectrometry) were run using primary antibodies against these proteins. Ultimately, in order for bands of interest to be visible, high concentrations of serum were required to be run on the gel and protein load over-saturation and cross-reactivity were seen across the blot. Due to the unreliability of over-saturated gel to correctly resolve proteins by molecular weight, these proteins were not taken forward for further analysis by Western blot. In general, proteins with a mean normalised abundance of <100,000 in mass spectrometry were difficult to visualise reliably by WB without sample fractionation.

4.3.2. Leptospirosis

During mass spectrometry analysis of serum proteome, leptospirosis positive sera were seen to be highly distinct from other sera groups. A total of 108 proteins were found to be in significantly different abundance when comparing leptospirosis positive sera to all other sera groups. In order to further demonstrate the ability of label-free quantitative mass spectrometry to identify novel biomarkers for infectious disease, two proteins of interest, serum amyloid A1 (SAA1) and carbonic anhydrase III (CAH3) were taken forward for further analysis.

4.3.2.1. Serum amyloid A1 ELISA

Quantitative mass spectrometry identified the protein serum amyloid A1 (SAA1) as increased in abundance in leptospirosis positive sera compared to all other groups (Log^2 Fold Change =5.942, p=0.033). **Figure 4.6** shows the normalised abundance data for SAA1 as measured by mass spectrometry.

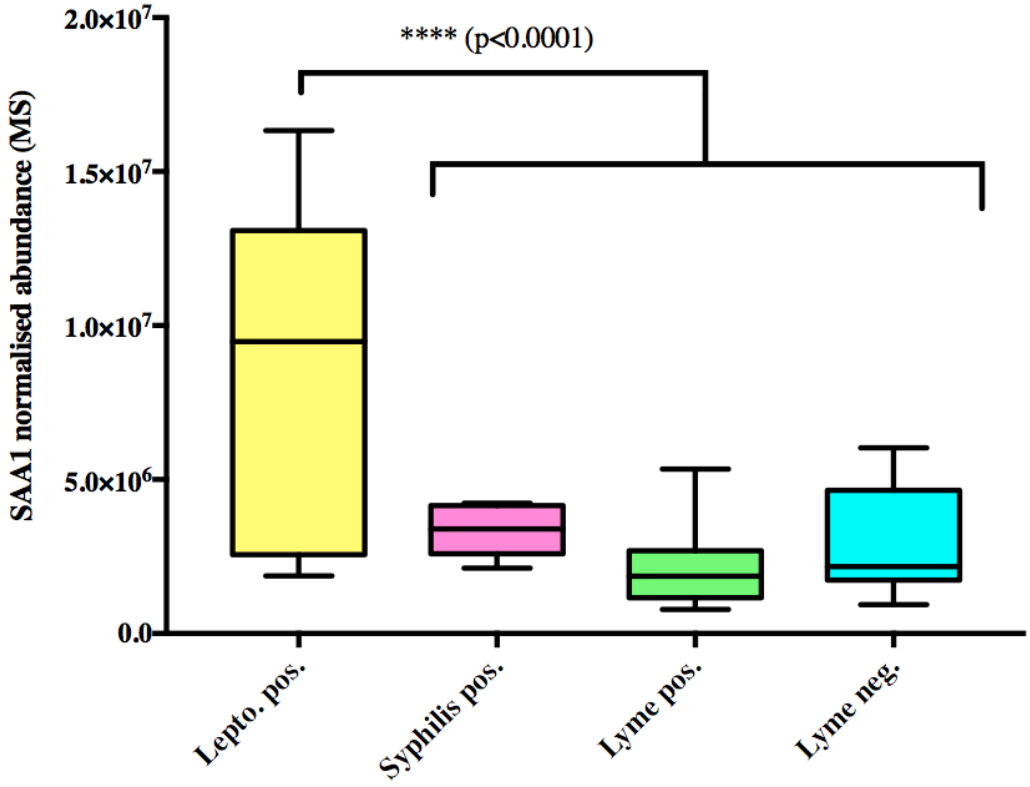


Figure 4.6: Chart showing normalised protein abundance data for serum amyloid A1 (SAA1) in human sera by label-free quantitative mass spectrometry. Error bars represent standard error of the mean (SEM). Sample groups; Leptospirosis positive n=5, syphilis positive n=5, Lyme disease seropositive (by STT) n=15, Seronegative (by STT) n=15.

SAA1 is a commonly measured protein in serum and several pre-optimised ELISA kits are commercially available for quantification in human sera. As with LCN2, a Duoset ELISA kit (R&D Human-SAA1 DuoSet ELISA kit DY3019-05) was chosen to allow processing of a large number of samples across multiple plates. In general, kit protocol was followed as described; however, to optimise the ELISA for use with serum samples, a diluent containing PBS and 10% bovine serum albumin was used (Materials and Methods: 3.3.2). **Figure 4.7** shows a graph of SAA1 serum concentration, as measured by ELISA. **Figure 4.7(A)** shows all sample groups separately, with each point representing one serum sample. The mean concentration for each group were as follows: UK Lyme disease positive = 7592(±1635) ng/mL (n=26), UK Lyme disease seronegative = 7070(±2347) ng/mL (n=17), Czech Lyme disease positive = 7864(±2109) ng/mL (n=25), Czech Lyme disease seronegative = 7071(±17.51)ng/mL (n=24), leptospirosis positive = 17922(±2311)ng/mL (n=18), syphilis positive =12386(±974)ng/mL n=10, NHD = 6039(±1136)ng/mL (n=30). **Figure 4.7(B)** shows a graph of the ELISA data split into leptospirosis, NHD and a final group with all other sera. The mean concentration of leptospirosis positive sera was 15990(±880.4)ng/mL, while in the related-control group the mean was 8044(±466.5)ng/mL. This equates to a log[2]fold change increase of 0.991 in leptospirosis positive. The data demonstrate the ability of mass spectrometry to identify potential biomarkers in human sera even in this small sample set (leptospirosis positive n=5).

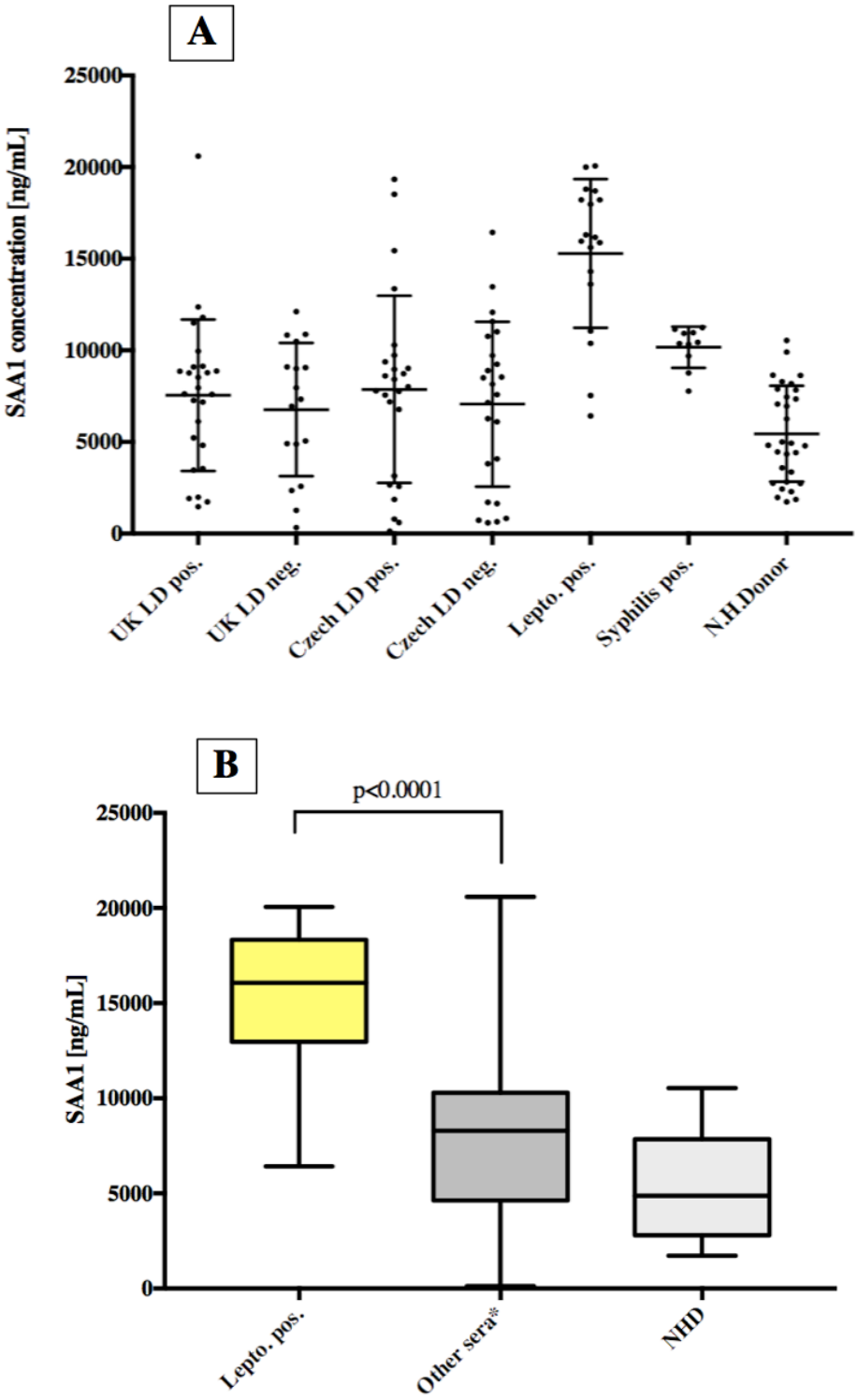


Figure 4.7: Graphs showing serum amyloid A1 abundance as measured by ELISA (R&D Human-SAA1 DuoSet ELISA kit DY3019-05). A: Abundance data for all groups. B: Abundance data for combined Lyme disease positive groups (UK and Czech Lyme disease seropositive) and combined Lyme disease seronegative groups (UK and Czech Seronegative) along with the NHD group. Combined Lyme disease positive n=52, combined Lyme disease seronegative n=50, N.H donor n=30

4.3.2.3 Carbonic anhydrase III ELISA

Quantitative mass spectrometry identified the protein carbonic anhydrase III (CAH3) as highly increased in abundance in leptospirosis positive sera compared to all other groups (Log^2 Fold Change =7.327, p=0.006).. **Figure 4.8** shows the normalised abundance data for CAH3 as measured by mass spectrometry.

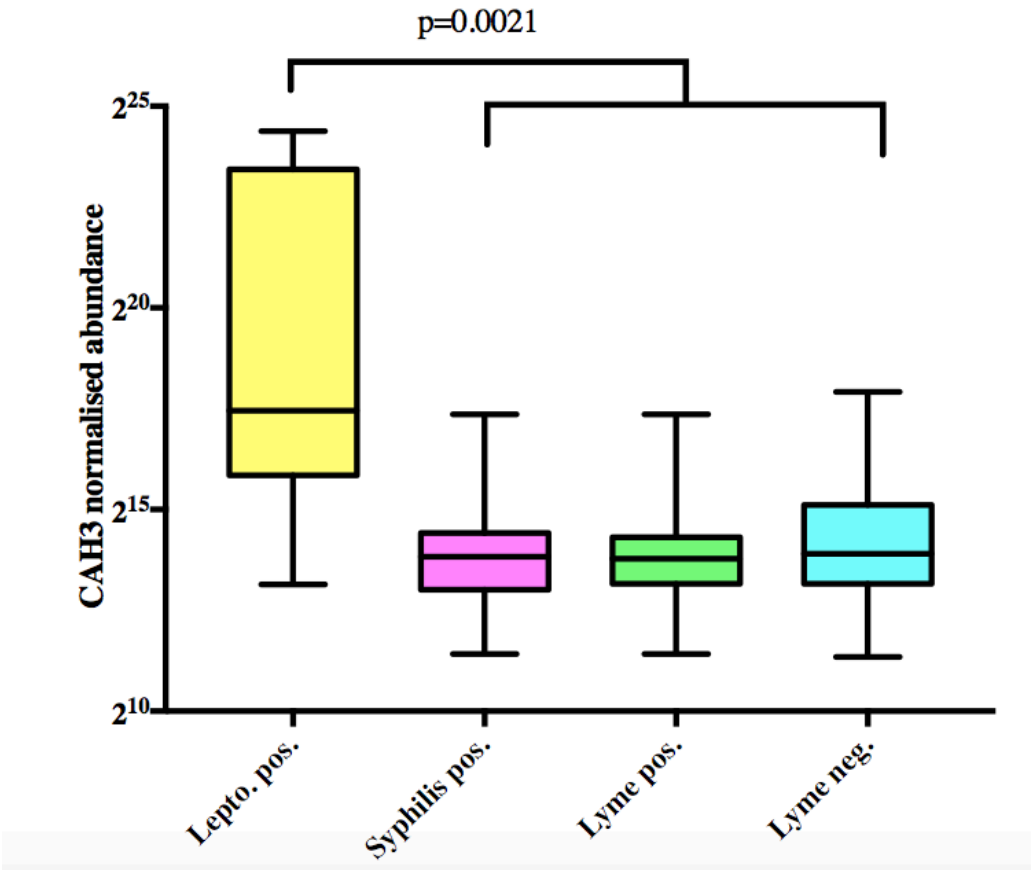


Figure 4.8: Chart showing normalised protein abundance data for carbonic anhydrase III (CAH3) in human sera by label-free quantitative mass spectrometry. Error bars represent standard error of the mean (SEM). Sample groups; Leptospirosis positive n=5, syphilis positive n=5, Lyme disease seropositive (by STT) n=15, Seronegative (by STT) n=15.

In the absence of commercial ELISA kits for large-scale analysis of sample (>39 samples), an ELISA platform for CAH3 was developed using a primary and secondary antibody pair for the protein (Novus CAH3 antibody pair H00000761-AP11)(Materials and Methods: 2.3.2).. In order to optimise the concentration of primary and secondary antibody for ELISA, a checkerboard antibody titration assay was undertaken. **Figure 4.9(A)** shows the layout of the 96-well plate for titration assay. **Figure 4.9(B)** shows the results of this assay, given as signal to noise ratios for each possible capture and detection antibody detection. A higher signal to noise ratio is desirable. An antibody at a concentration of 2μg/mL and detection antibody at a concentration of 0.5μg/mL (signal/noise=20.072) were chosen as suitable concentrations for the CAH3 ELISA.

A

	1	2	3	4	5	6	7	8	9	10	11	12
	1ug/mL capture	1ug/mL capture	2ug/mL capture	2ug/mL capture	4ug/mL capture	4ug/mL capture	1ug/mL capture	1ug/mL capture	2ug/mL capture	2ug/mL capture	4ug/mL capture	4ug/mL capture
A	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ
B	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard
C	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard
D	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard
E	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ	Δ
F	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard	10ng/mL standard
G	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard	20ng/mL standard
H	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard	40ng/mL standard

0.25ug/mL detection
0.5ug/mL detection
1 ug/mL detection
2 ug/mL detection
Δ Diluent only, no standard

B

Capture	Detection 0.25 $\mu g/mL$	0.5 $\mu g/mL$	1 $\mu g/mL$	2$\mu g/mL$
1 $\mu g/mL$	16.153	18.5135	16.981	15.705
2 $\mu g/mL$	16.014	20.072	18.201	15.84
4 $\mu g/mL$	15.237	16.531	14.714	16.147

Figure 4.9: Optimisation of primary and secondary antibodies for a CAH3 ELISA. A: Table showing the 96-well plate layout of checkerboard optimisation assay. Standards of CAH3 at the following dilutions were used to assess range: 10ng/mL, 20ng/mL and 40ng/mL. Δ = wells with only sample diluent (no standard). B: Signal-to-noise ratio calculated by dividing OD read for each sample by the corresponding OD read for negative control (Δ = diluent only). This was performed for each standard and a mean signal-to-noise ratio value was calculated for each primary/secondary antibody pairing.

In order to quantify CAH3 abundance in serum samples, a standard curve was developed using a commercial recombinant purified CAH3 protein (abcam ab82823). Following the suggested total range of detection for the antibody pair used (1-100ng/mL), standards at 1,5,10,20,40,50,80 and 100ng/mL CAH3 concentration were used. **Figure 4.10** shows the standard curve generated using these standard dilutions.

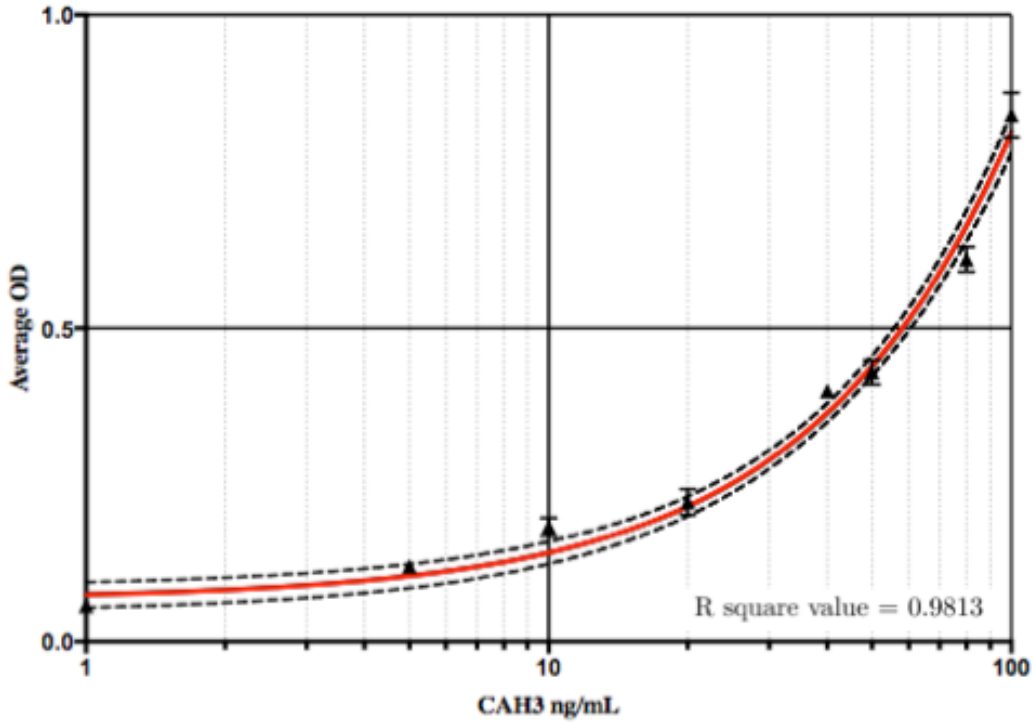


Figure 4.10: A standard curve for carbonic anhydrase III protein using recombinant protein. Standards were run in triplicate, with error bars representing standard error of the mean (SEM). The red shows and example of the standard curve (R-squared value=0.9813) used to quantify CAH3 abundance in serum samples (plate dependent).

A total of 103 serum samples were tested for CAH3 abundance using the optimised ELISA protocol. **Figure 4.12(A)** shows all sample groups separately, with each point representing one serum sample. The mean concentration for each group were as follows: UK Lyme disease positive = 117.8(±65.6)ng/mL (n=26), UK Lyme disease seronegative = 138(±68.1)ng/mL (n=17), Czech Lyme disease positive = 139.8(±57) ng/mL (n=25), Czech Lyme disease seronegative = 92.08(±35.52)ng/mL (n=24), leptospirosis positive = 290.5(±98.8) ng/mL (n=18), syphilis positive =145.8(±28.2)ng/mL n=10, NHD = 104.8(±38)ng/mL (n=30). **Figure 4.7(B)** shows a graph of the ELISA data split into leptospirosis, NHD and a final group with all other sera. The mean concentration of the combined other sera group was 127.8(±12.13)ng/mL. This equates to a log[2]fold change increase of 1.185 in leptospirosis positive. This data corroborates the mass spectrometry findings with a large increase in abundance of CAH3 in leptospirosis positive sera, with other sera groups found to be in a similar range to NHD controls.

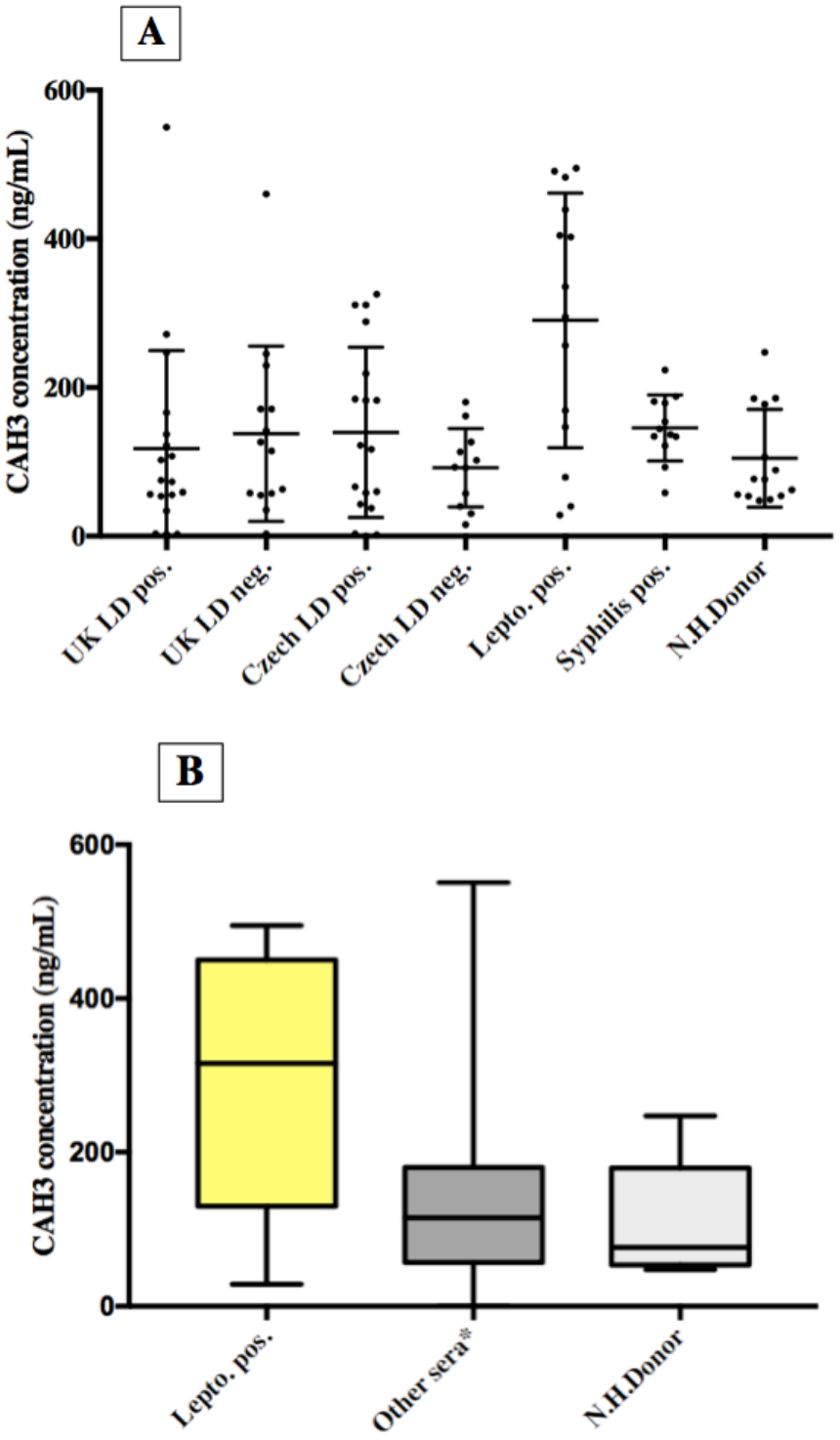


Figure 4.11: Graphs showing CAH3 abundance as measured by ELISA. A: Abundance data for all groups. B: Abundance data for combined Lyme disease positive groups (UK and Czech Lyme disease seropositive) and combined Lyme disease seronegative groups (UK and Czech Seronegative) along with the NHD group. Combined Lyme disease positive n=52, combined Lyme disease seronegative n=50, N.H donor n=30.

4.4. DISCUSSION

In chapter 4, label-free quantitative mass spectrometry revealed that the serum proteome of Lyme disease seropositive patients was largely similar to that of those that had been tested for Lyme by STT testing but were found to be seronegative. Despite this, several proteins were found to be differentially abundant between these groups including lipocalin-2, which was increased in Lyme disease seropositive sera, and tubulin-alpha 4A, which was seen to be decreased in Lyme disease seropositive sera. WB and ELISA methods were able to corroborate lipocalin-2 findings, with the protein found to be in significantly higher abundance in Lyme disease seropositive (99.93(±10.04)ng/mL than Seronegative 63.08(±6.463)ng/mL. WB analysis did not identify any significant differences for tubulin alpha abundance between these groups. To further assess the ability of mass spectrometry to identify potential biomarkers the proteins serum amyloid A1 (SAA1) and carbonic anhydrase III (CAH3), previously identified to be highly increased in abundance in leptospirosis positive sera. ELISA data for SAA1 and CAH3 corresponded to mass spectrometry findings, with leptospirosis sera having the highest abundance of these proteins (17922(±2311)ng/mL and 290.5(±98.8)ng/mL respectively).

A recurring problem during Western blotting for proteins of interest was the presence of smeared bands and/or the presence of unspecific bands. For several proteins, these problems reoccurred despite using alternative primary antibodies. There are several possible reasons for the difficulty in developing clear blots with a single clear band at the expected molecular weight. The presence of smeared bands is often an indication of protein degradation, with smaller fragments of protein that contain the target antigen spread below the band of interest. As disruption of cell membranes during

blood collection and subsequent processing to serum releases proteases, kinases and phosphatases, degradation of sample begins early (239). While enzyme activity is halted when samples are frozen, once thawing begins samples are again susceptible to degradation. Another possible explanation for the presence of multiple bands is poor specificity in the primary or secondary antibody. In a highly complex biological sample such as total serum, the possibility of non-specific bands developing is greater. The use of monoclonal antibodies against a single antigen theoretically reduces the possibility of non-specific bands forming when compared to polyclonal antibodies produced by multiple B-cell lineages. Despite this, non-specific binding was still observed when using monoclonal primary antibodies for several proteins. Poor WB band formation may be due to the complex nature of serum, several orders of magnitude exist between abundance of several of the proteins-of-interest identified by mass spectrometry and the most abundant proteins in serum. When attempting to blot for lower abundance proteins, larger total protein loads per well are required to allow detection. This may have resulted in separation of proteins as the gel is overwhelmed by protein load. Budgetary constraints meant that samples not used in mass spectrometry did not undergo depletion by Pierce top-12 columns. In future work, sample fractionation by other means such as simple pore filters or more advanced 96-well plates that fractionate proteins based on isoelectric point could be used. In the context of this study, the use of non-fractionated serum better represents the biological specimen that would likely be used in any potential biomarker-based diagnostic assay.

Mass spectrometry data indicated that tubulin alpha 4A and several actin isoforms were differentially abundant across sample groups. This finding was unusual as these proteins are generally considered to be found at consistent levels regardless of

physiological state, and both tubulin alpha and actin alpha are recommended as "housekeeping" loading controls for Western blot. In this study, alpha tubulin was found to be consistent across Lyme disease seropositive and Seronegative by WB at the resolution possible. In previous studies, however, these proteins have been shown to have inconsistent abundance across sample types(240-242). For this reason, a protocol using total protein staining by SYPRO ruby was used instead of a traditional housekeeping gene during WB analyses for all proteins.

The serum acute phase-response protein, serum amyloid A1 (SAA1), was previously found to be increased in abundance in leptospirosis positive sera by mass spectrometry (Chapter 3) (Log[2]Fold Change=5.942, p=5.0E-05) and was again found to be increased in leptospirosis positive sera by ELISA (leptospirosis positive: 17922(±2311)ng/mL, other sera 8106(±1045)). Acute-phase serum amyloid A proteins (A-SAAs) are secreted during inflammation acute phase and have a variety of functions including immune cell recruitment to inflammatory sites. Differences in abundance of SAA1 have previously been identified in a 2017 study by Tan et al when comparing mild and severe leptospirosis cases using 2DE analysis LC-MS/MS. In that study both serum amyloid A-1 and A-2 were significantly increased between healthy to mild and again between mild to severe leptospirosis cases (243). A-SAAs have been shown to be deposited in the kidney and liver during infection and probably contribute to kidney and liver failure commonly reported in severe leptospirosis (244).

For LCN2 and SAA1, commercial ELISA kits were able to corroborate the findings found by mass spectrometry. The LCN2 ELISA showed excellent performance (correlation) for the development of standard curves (LCN2 R squared value =

0.9901. For SAA1, correlation was found to be poor at highest levels of standards with very little increase in OD seen between the top two standards (50ng/mL and 100ng/mL). Despite this, correlation was very good at lower standard concentrations, and the majority of serum samples tested were well within this range. In future, a better standard range and/or further optimisation of primary and secondary antibodies could improve the precision of SAA1 levels as measured by this ELISA. Following development and optimisation, the CAH3 ELISA was able to quantify CAH3 protein across serum samples. In order to determine if protein detection was affected by standard curve diluent and biological sample (serum), a spike-and-recovery and linearity-of-dilution assessment for ELISA could be undertaken.

4.4.1. Conclusions

Several proteins found to be differentially abundant in different disease states by mass spectrometry in chapter 4 were successfully quantified by WB and/or ELISA methods. In general, results for WB and ELISA corroborated mass spectrometry findings. Lipocalin 2 protein was shown to be significantly increased in abundance in Lyme disease seropositive over seronegative by WB and ELISA. The results of this study demonstrated the ability of quantitative mass spectrometry to identify proteins that are differentially abundant in the serum of Lyme disease and other infectious-disease patients.

CHAPTER 5: RNA-SEQ ANALYSIS OF WHOLE BLOOD FROM PATIENTS WITH LYME DISEASE

5.1. INTRODUCTION

The impact of Lyme disease on the host can result in changes in protein abundance and potentially post-translational regulation. However, often at the root of these signalling events are changes in the host transcriptome. Blood represents an excellent sentinel to examine the impact of a pathogen on the host as this sample contains multiple cell types whose abundance and transcriptomic profile reflects pathogen/host interactions. To date, relatively few transcriptomic-based studies have been published on Lyme disease. Of these, two have focused on studying gene expression in the *Borrelia* pathogen itself (245, 246) both *in vitro* and *in vivo*. A 2016 study by Bouquet *et al* examined the molecular basis of acute Lyme and ensuing post-treament symptoms by undertaking a longitudinal study of patients followed up to 6 months post treatment. They found a large change in gene expression during acute infection showing fewer than 50% differentially expressed genes in common with other infectious diseases. No differential gene expression signature was observed between Lyme disease patients with resolved illness to those with persistent symptoms (247)(Bouquet 2016). In this study, 14 whole-blood samples from 8 Lyme disease patients underwent RNA extraction and subsequent sequencing by Illumina RNA-seq. Sequencing data from previous studies were used to form control groups and allow relative gene expression to be calculated. Machine learning algorithms based on DGE were used to identify a panel of genes that could distinguish early Lyme disease patient samples from those of the control group. Analysis of gene expression also allowed study of the host response to Lyme disease and the identification of

associated biological pathways. Finally, RT-PCR was used to corroborate findings in a different platform and showed significant changes in presence of gene transcripts in genes associated with eIF2a signalling, as identified by RNA-seq and downstream analyses.

RNA sequencing (RNA-seq) is a technique that utilises Next Generation Sequencing (NGS) to sequence fragments of RNA in a biological sample that can be mapped to expressed RNA in the host genome (248). The technology allows quantification of gene expression at the time of sampling, and can therefore give a representation of host biological processes occurring during a particular biological state including acute infection, progression of a disease and response to therapy. RNA-seq protocols involve the conversion of RNA to cDNA fragments which are subjected to NGS, producing short sequences that correspond to a specific host gene. Single-read sequencing methods involve reading of the cDNA from one end only. Paired-end methods, including that used in this study, involves the sequencing of cDNA from both ends of the fragment (249). While more time consuming and costly, paired-end reading allows more accurate mapping of fragments relative to the genome and allows detection of insertion-deletion variants. RNA-seq produces raw sequence data, often provided in text-based FASTQ files. The data provides only nucleotide sequences together with a corresponding quality score for each base, a measure of the probability of correct base calling based on properties including signal-to-noise ratios. The data are commonly filtered based on quality score, and then sequences can be mapped to their corresponding position on the host genome using mapping software. The number of reads mapped to certain genes allows a quantification of the expression of that gene in that sample. Comparison of two or more cohorts allow Differential Gene

Expression (DGE) to be identified between groups. This is particularly important in biomarker discovery studies, where identifying quantifiable changes in gene expression in a certain group may be of more value than understanding underlying biological processes.

Several other tools exist for the use of RNA sequencing data, either from the raw read data itself or from DGE output such as DESeq2. These include pathway analysis tools such as Ingenuity Pathway Analysis (IPA), software that estimates differential expression of biological pathway based on the increased or decreased expression of specific combinations of genes in the cohorts compared (250). Similar programmes are available that predict the make-up and quantity of immune cells present in a biological sample or group of samples based on gene expression data (251).

The sheer abundance of data generated by RNA sequencing makes it an attractive target for the discovery of biological biomarkers with diagnostic or prognostic properties. The availability of expression data for tens of thousands of genes allows for identification of genes and gene panels that are diagnostic of a certain condition, often with a high degree of accuracy; however, whole RNA sequencing of biological samples is expensive and time consuming, and it is unlikely that routine diagnostic tests could be based on sequencing of patient sample in the near future. Diagnostic tests could theoretically be based on performing multiplex RT-PCR assays for the detection of a panel of host genes, and this has been proposed before, including for distinguishing between viral and bacterial respiratory infections (252). RT-PCR methods for host-genes have also been used for over a decade in quantification of minimal residual disease, the presence of tumor mass remaining following treatment

in leukemia patients (253). RT-PCR methods for the detection of infectious pathogen genome or genes present in the host are far more commonly used as diagnostics, with the presence of high levels of pathogen nucleic acid often being a clear indication of ongoing infection. In the case of Lyme disease, previous studies attempting to use PCR for the detection of *Borrelia* have been largely unsuccessful, or have required the use of biological samples such as erythema migrans skin biopsies that are not appropriate for routine diagnosis (254, 255). In samples commonly used for diagnostics, including blood and urine, *Borrelia* are often present in too low abundance to be quantified using PCR, even during acute disease. In these cases, the value of RNA-seq often lies in the identification of genes that are highly differentially expressed between groups, or in the use of machine-learning algorithms to identify panels of genes that can distinguish a sample from a control group. RNA-seq together with downstream bioinformatics can also give a picture of the on-going host biological processes that underlie or influence disease progression. It can also identify certain biological pathways associated with a disease that have previously gone unrecognised.

5.2. MATERIALS AND METHODS

Table 5.1 shows summary data of samples used during RNA analyses. Group 1: Lyme disease patients, were provided as whole-blood samples in PAXgene tubes. They were sequenced and mapped as described in the main Materials and Methods chapter (3.7.4). Group 2: Controls – Ebola virus (EBOV) positive and convalescent, were provided as trimmed Illumina sequencing data. These data were processed using the same pipeline as the Group 1. Group 3: Controls – NHD, were obtained as trimmed Illumina sequencing data (online database). Again, these data were processed as with groups 1 and 2.

Group 1: Lyme patients

Patient ID	Sex	Sample ID	Sample date	Days post presentation	CLIA IgG	CLIA IgM	WB IgG	WB IgM
L 1	M	**1/84**	16/10/17	84	Neg	Bor	Neg	Neg
L 2	M	**2/92**	30/10/17	92	Pos	Neg	Pos	Neg
L 7	F	**7/17**	06/09/17	17	Pos	Pos	Pos	Neg
		7/30	19/09/17	30	Pos	Neg	Bor	Neg
		7/88	16/11/17	88	Pos	Neg	Pos	Neg
L 8	M	**8/1**	05/09/17	1	Neg	Neg	x	x
		8/20	25/09/17	20	Neg	Neg	x	x
L 9	M	**9/18**	12/09/17	18	Neg	Pos	Neg	Pos
		9/33	27/09/17	33	Neg	Pos	Neg	Pos
L 10	M	**10/1**	18/09/17	1	Bor	Pos	Bor	Pos
L 11	F	**11/1**	02/10/17	1	Pos	Pos	Neg	Pos
		11/14	16/10/17	14	Pos	Pos	Neg	Pos
L 12	F	**12/1**	09/11/17	1	Pos	Pos	Bor	Pos
		12/18	27/11/17	18	Pos	Pos	Bor	Pos

Peripheral venous whole-blood samples from patients with acute or late/convalescent Lyme disease at time of collection. All samples provided by Ceske Budjovice Hosptial, Czech Republic.

Group 2: Controls - EBOV

No.	Sample ID	Sex	Age	Collection date	EBOV PCR
1	**E-51**	M	40	x/10/19	Neg
2	**E-26**	F	26	x/10/19	Neg
3	**E-31**	M	30	x/10/19	Neg
4	**E-12**	M	28	x/10/19	Neg
5	**E-46**	F	36	x/10/19	Neg
6	**E-38**	F	41	x/10/19	Neg
7	**E-61**	M	21	x/10/19	Neg
8	**E-42**	M	22	x/10/19	Neg
9	**E-50**	F	28	x/10/19	Neg
10	**E-55**	F	30	x/10/19	Neg
11	**EP-28**	M	19	x/10/19	Pos
12	**EP-12**	M	28	x/10/19	Pos
13	**EP-14**	F	27	x/10/19	Pos
14	**EP-33**	F	56	x/10/19	Pos
15	**EP-19**	M	43	x/10/19	Pos
16	**EP-20**	M	25	x/10/19	Pos

RNA sequencing data from peripheral venous blood. Illumina sequencing data provided by the European Mobile Laboratory Consortium, Hamburg, Germany. All samples collected during West African Ebola outbreak. Samples 1-10 are from recovered/convalescent individuals (PCR negative). 11-16 are from acute EBOV patients (PCR positive).

Group 3: Controls - Normal healthy donors

Sample ID	Sex	Collection date
H1	M	x/x/2014
H2	M	x/x/2015
H3	M	x/x/2016
H4	F	x/x/2017
H5	F	x/x/2018
H6	F	x/x/2019

Peripheral venous blood from normal healthy donors, collected in PAXgene tubes. Taken as part of a study by Shin et al. RNA sequencing data.

Table 5.1: Summary data for Lyme disease samples used in this study together with details of control groups. All Lyme disease samples were sequenced for use in this study, while control groups were based on sequence data from previous studies (not authors own work (256, 257))

5.3. RESULTS

5.3.1. RNA sequencing (RNA-seq)

Samples were collected in PAXgene blood RNA tubes allowing for immediate stabilisation of intracellular RNA during blood sample collection. A PAXgene purification kit was then used for the purification of intracellular RNA from whole blood. To further purify samples, a Qiagen spin-column based kit was used and samples were run on a denaturing agarose gel to assess their quality by presence of distinct 28S and 18S rRNA bands (Materials and Methods 2.7.1). Samples were then run using Illumina RNA-seq by the Centre for Genomic Research (CGR), University of Liverpool (Materials and Methods 2.7.4). Sequencing statistics for each sample are provided in **table 5.2.** A mean of 78.42 (±52.5) million reads were sequenced per sample. This is considered very good for blood samples and provided ample depth for downstream DGE analyses. Following sequencing, trimmed reads provided by CGR were mapped and quantified following the method described in Materials and Methods 2.7.4. Briefly, trimmed read data for forward and reverse sequencing were concatenated, aligned to the human genome using TopHat and were then converted to read count files using featureCounts. This pathway was also used for raw read data used in the Ebola positive and NHD (normal healthy donor) control groups. Figure 5.1 shows a Venn diagram of all human genes detected per sample group. A total of 10842 genes were detected in all sample groups compared.

	ID	No. of reads (million)	Reverse read length mean (bp.)	Forward read length mean (bp.)	Mean base quality score	A content %	C content %	G content %	T content %
1	L 1/84	60.62	142.5	142.7	38.9	23.2	25.0	26.2	25.6
2	L 2/92	51.01	143.4	143.3	37.9	24.5	24.1	24.7	26.6
3	L 7/17	62.19	143.6	144.2	37.7	24.3	24.4	26.3	25.0
4	L 7/30	67.82	143.2	143.6	38.1	24.6	23.7	25.1	26.6
5	L 7/88	64.02	142.6	143.4	38.5	23.9	24.6	24.9	26.6
6	L 8/1	64.02	142.7	143.4	38.4	23.8	24.5	24.8	26.9
7	L 8/20	55.32	143.5	144.5	37.9	24.7	23.7	25.4	26.2
8	L 9/18	71.21	146.1	145.5	39.1	24.3	24.6	26.1	25.0
9	L 9/33	120.54	144.2	143.4	38.4	24.4	24.6	24.9	26.1
10	L 10/1	251.11	147.3	144.5	38.1	23.7	24.7	25.8	25.8
11	L 11/1	62.12	146.5	145.5	37.2	23.6	24.6	25.1	26.7
12	L 11/14	54.12	143.2	142.9	38.2	24.8	24.8	24.9	25.5
13	L 12/1	59.83	144.4	144.3	38.8	23.9	24.7	25.9	25.5
14	L 12/18	53.88	144.6	145.2	37.9	24.3	24.6	26.1	25.0

Table 5.2: RNA sequencing statistics for the 14 whole-blood samples included in this study

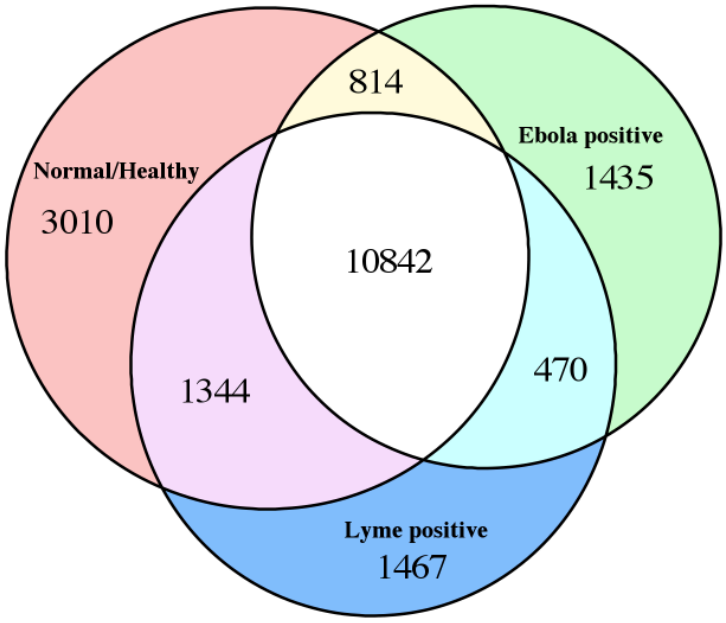


Figure 5.1: Venn diagram showing the number of unique human genes successfully detected and quantified following RNA sequencing and genome mapping for each of the three groups.

5.3.2. Differential gene expression: Lyme disease

To understand differences in gene expression in patients with Lyme disease, a differential gene expression study was undertaken, comparing 6 early/mid Lyme disease patient sample data (Samples: L 8/1, L 10/1, L 11/1, L 12/1, L 11/14) against Ebola and NHD controls. All Lyme disease samples were positive or borderline for Lyme by CLIA IgG, CLIA IgM and IgM WB at the time of sample collection. Differential gene expression analysis was performed using a DESeq2 pathway, as described in Materials and Methods 2.7.4. Briefly, featureCount data was normalised by internal normalisation by DESeq2 and differential gene expression was calculated by use of negative binomial generalised linear models. DESEq2 results files provided a list of differentially expressed genes and a calculated log [2] fold change estimate between groups for each gene. Only genes found to be differentially expressed with a

statistical significance of $p<0.05$ are included in these data. In this study, 5 early/mid Lyme disease samples were compared to a control group of 13 control samples (composed of Ebola convalescent and NHD). **Figure 5.2** shows a MA plot of the DGE data. A total of 9435 genes were found to be significantly increased and 3817 decreased in the Lyme disease group.

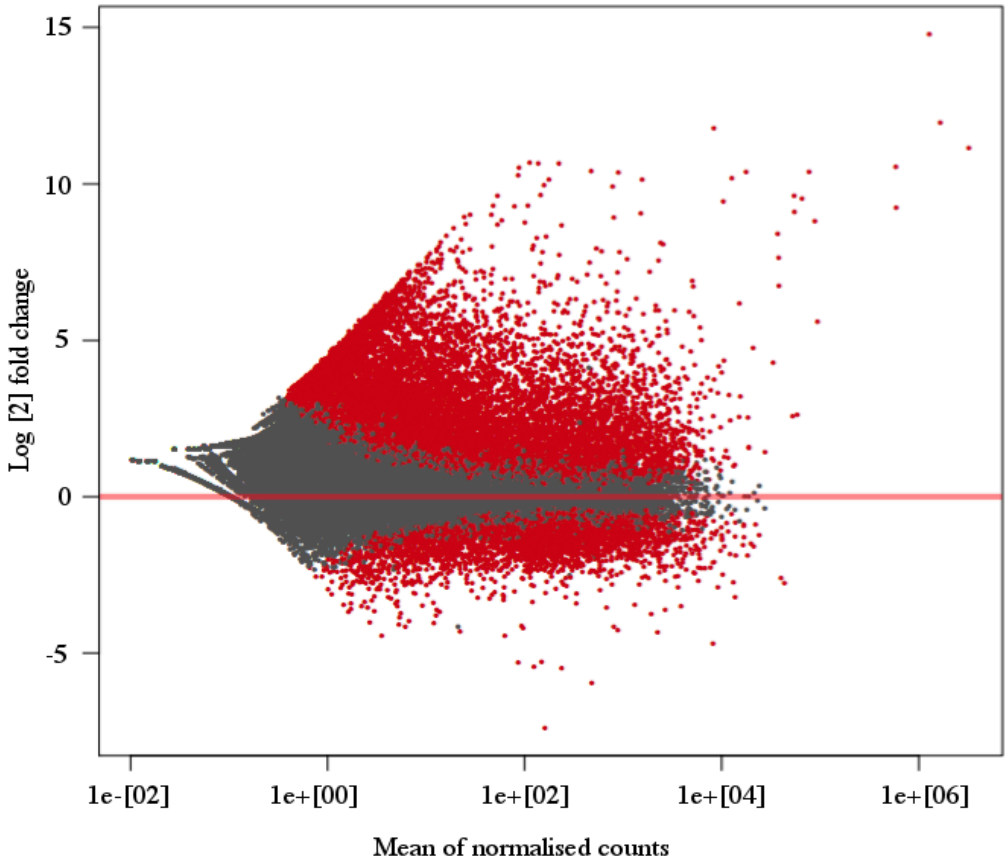


Figure 5.2: A MA-plot of DGE data from a comparison of early Lyme disease patient blood RNA to a combined infectious and normal/healthy control group. The data is transformed onto the M (log ratio) and A (mean average) scales and plotted. This allows visualisation of differences in the dataset. All points represent one gene, with red points showing genes that are significantly changed between groups.

In order to identify a panel of genes that were able to distinguish between the Lyme disease group and controls, the MLSeq package in R was used to generate a Probabilistic Linear Discriminant Analysis (PLDA) machine learning protocol. Normalised abundance data generated during DESeq2 were used as an input. Sample data were then split into training and test sets. This allowed development of the machine learning protocol and subsequent testing of the model to verify its accuracy. Developing a protocol using only a training set is likely to output genes that are indicative of the input sample data, rather than changes that are conserved across samples of that group. A PLDA (probabilistic linear discriminant analysis) based machine learning algorithm was used. From the training set, 9 genes were identified as discriminatory between Lyme and control groups. Of these, 6 had a higher mean read count value in the Lyme disease group (S100A9, EEF1A1, CSF3R, MYO1F, LITAF and CORO1A), while 3 (IFITM2, SELL and ACTB) and had a lower value in the Lyme disease group. Figure 5.3 shows the normalised read counts (Log [2]) of all samples within these groups. Using the machine learning testing set, values for the accuracy of the gene-panel (based on cut-off values) to discriminate between Lyme disease samples and controls were generated. Accuracy, sensitivity and specificity values nearing 100% were predicted based on the model. It should be noted, however, that these values were based on a very small number of samples being tested by the model, and without robust downstream analysis using additional samples; no meaningful conclusions about the model accuracy can be made at this time. The study limitations in regard to machine learning are discussed later in this chapter.

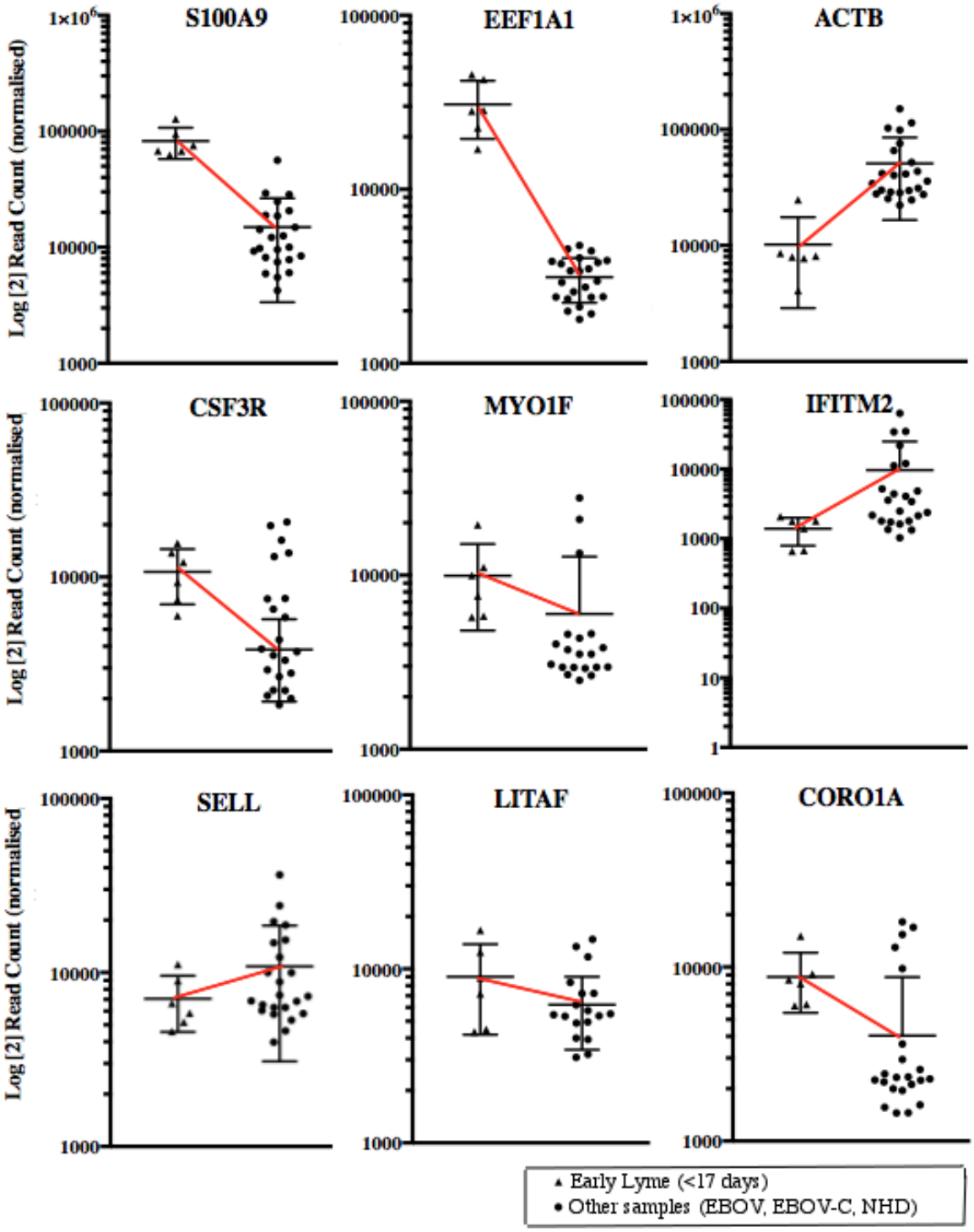


Figure 5.3: Graphs showing read count data for the 9 genes identified by PLDA-based machine learning as discriminatory between Lyme disease blood RNA and controls. Read counts for each gene (Log [2]) are plotted for each sample. Mean read count and SEM are shown as bars on the plots.

5.3.3. Differential gene expression: Early vs. Late Lyme disease

The Lyme disease samples used in this study were collected from patients with a range of days post-presentation, as described in the Materials and Methods section of this chapter. This allowed for a DGE study based on groupings of "early" and 'late/convalescent" Lyme disease samples with the aim of identifying markers that may be elevated in early disease and fall over time. A DESeq2 method was used comparing the early group (Sample: L 8/1, L 10/1, L 11/1, L 12/1) to late samples (L 1/84, L 2/92, L 7/88). A total of 727 genes were identified as differentially expressed between these groups. Of these, 501 were significantly decreased in late/convalescent samples, while 226 were significantly increased in this group. **Figure 5.4** shows a MA plot of this data following normalisation.

Following DESeq, a PLDA-based machine-learning algorithm was again applied to the data. This provided a list of 11 genes that were differentially expressed between the early and late group. To visualise these changes, the Lyme sample group was then split into three subgroups based on days post presentation; an "early" group (samples 8/1, 10/1, 11/1 and 12/1) a "mid" group (samples 7/17, 7/30, 8/20, 9/18, 9/33, 11/14 and 12/18) and a "late" group (samples 1/84, 2/92 and 7/88). Normalised read count data were then plotted for each gene across sample groups (**figure 5.5**). The DEseq2 protocol combined with PLDA-based machine-learning was able to identify a set of genes that fell during the course of Lyme disease. Proper matching of samples across time points was not possible due to lack of samples, the major limiting factor of this study. Generally, Lyme disease patients showed movement of blood transcriptome towards that found in the normal/healthy control group.

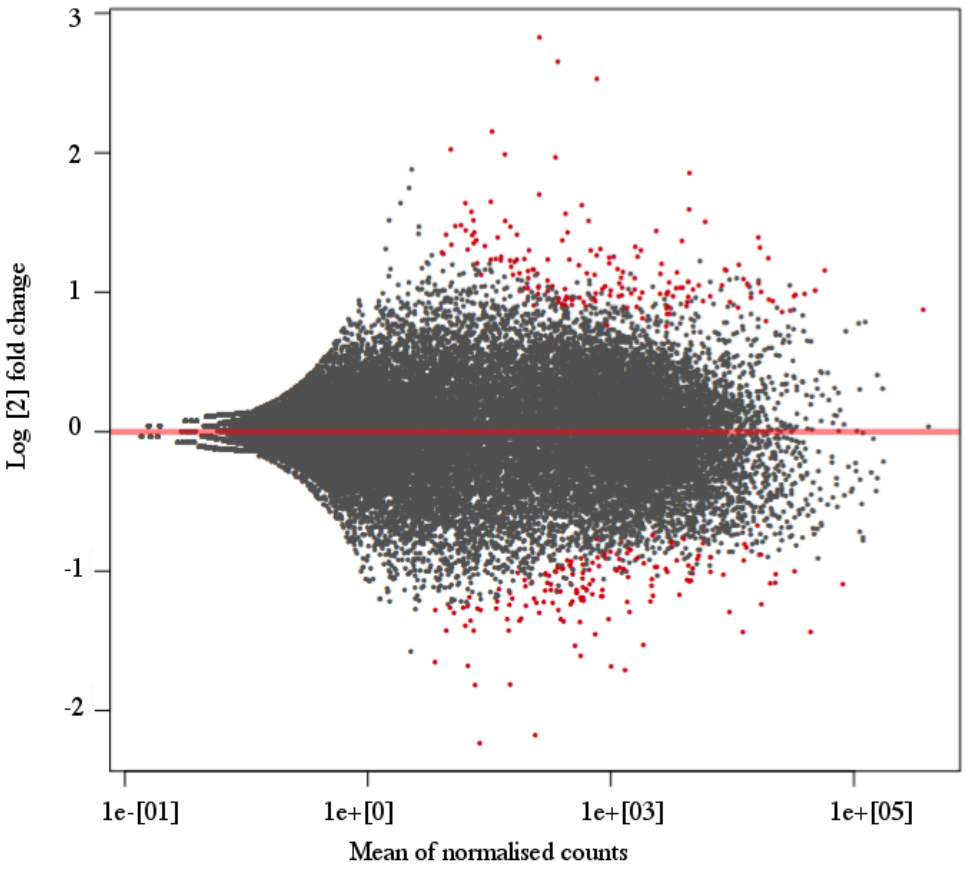


Figure 5.4: MA-plot of DGE data from early Lyme disease vs. mid/late Lyme disease as based on days post presentation. Each point represents a human gene, with points in red showing genes found to be significantly differentially expressed between groups.

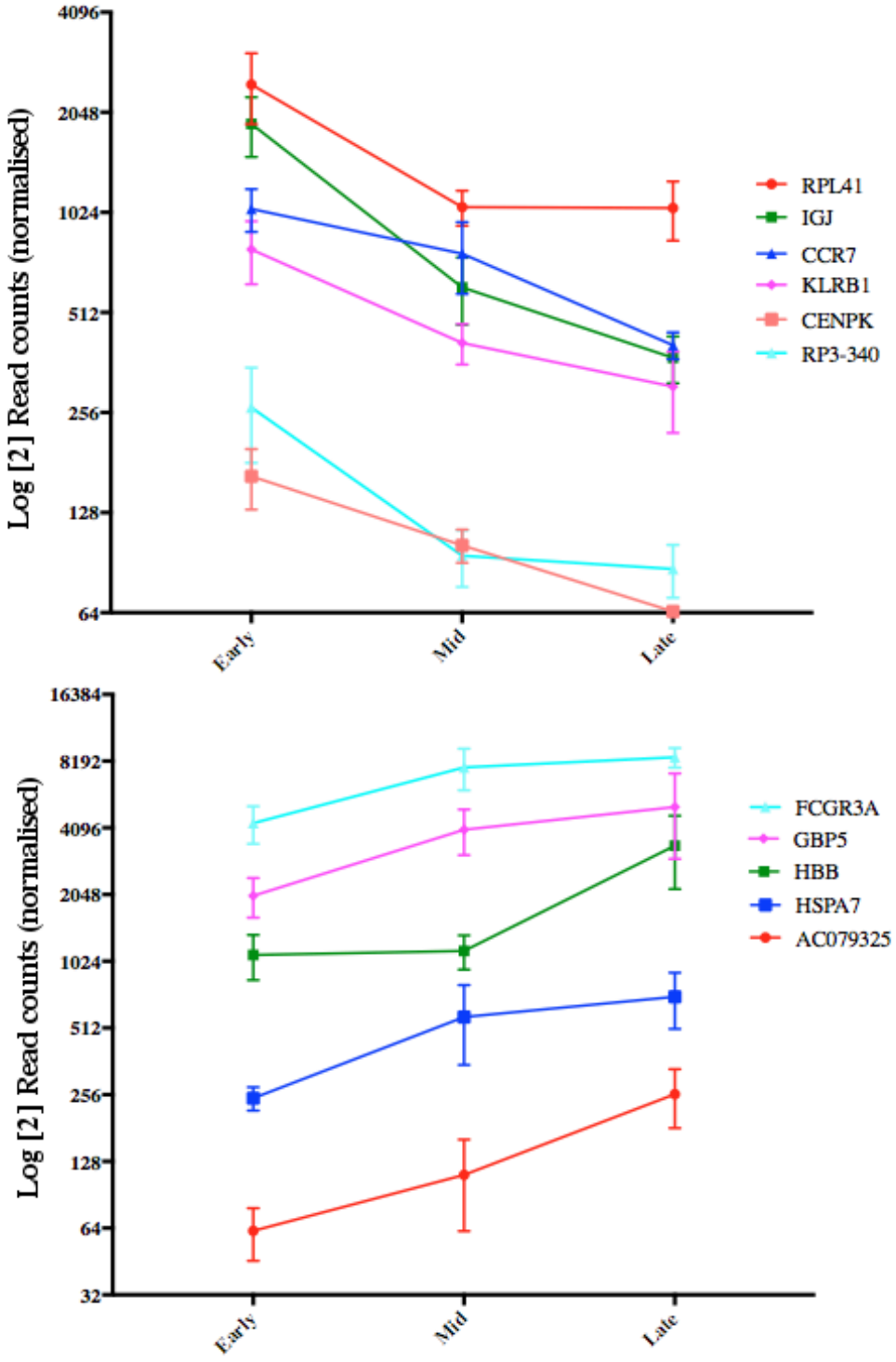


Figure 5.5: Graphs showing read counts (Log [2]) over time of 11 genes identified by PLDA-based machine learning as discriminatory between early and late Lyme disease Read count (Log [2]) data are shown for the 6 genes found to be decreased in the late group (top) and for the 5 genes found to be increased in the late group (bottom). Each point shows the mean value for that gene per group, with SEM error bars.

5.3.4 Ingenuity pathway analysis: Lyme disease

To gain a better understanding of the host response to Lyme disease, a further DGE experiment was conducted using early Lyme disease samples (Samples: L 8/1, L 10/1, L 11/1, L 12/1, L 11/14) and comparing to the healthy control group (Samples: H1-H6) using DESeq2. This generated a list of 6301 differentially expressed genes between these groups. This data was then input into the Ingenuity Pathway Analysis (IPA) software. IPA uses DGE data for pathway analyses based on differential expression of specific genes between groups. IPA generates a z-score and p-value for each pathway identified - based on the number of specific pathway genes identified and the differential expression of these between groups. A positive z-score indicates up regulation of that pathway, while a negative z-score indicates down regulation. For pathways in which no activity pattern is available, either through contradiction in expression of certain genes, or lack of reference material for the pathway, no z-score is given; in these instances a high p-value for the pathway indicates that the pathway is being influenced between group in some way, without a clear direction. **Figure 5.6** shows the top 15 pathways identified by IPA when comparing Lyme disease patients to healthy controls. Disruption of oxidative phosphorylation was identified as the top influenced pathway (based on p-value), showing an increase in expression of this pathway in Lyme disease patients. Further to this, IPA identified "mitochondrial dysfunction" as the top biological process in Lyme disease patients, when compared to healthy controls. The eIF2a signalling pathway was predicted to be highly increased in Lyme disease patients, along with regulation of the related initiation factor protein eIF4. The eIF2a pathway had previously been identified in previous transcriptomic-based studies of Lyme disease, where it was also found to be one of the top pathways influenced by Lyme disease (258). **Figure 5.7** shows a diagram generated by IPA for

the eIF2a signalling pathway. Interestingly, iron homeostasis signalling was found to be consistently disrupted in Lyme disease patients. The iron homeostasis associated protein lipocalin-2 has been described as a potential biomarker of Lyme in chapters 3 and 4. Pathways associated with cytoskeletal rearrangement, including regulation of actin-based motility by Rho were also identified by IPA. The potential association between host cell rearrangement disruption and Lyme disease has also been discussed in earlier chapters. The association between several of these biological processes and Lyme disease are unclear, but have consistently been identified as disrupted throughout this study.

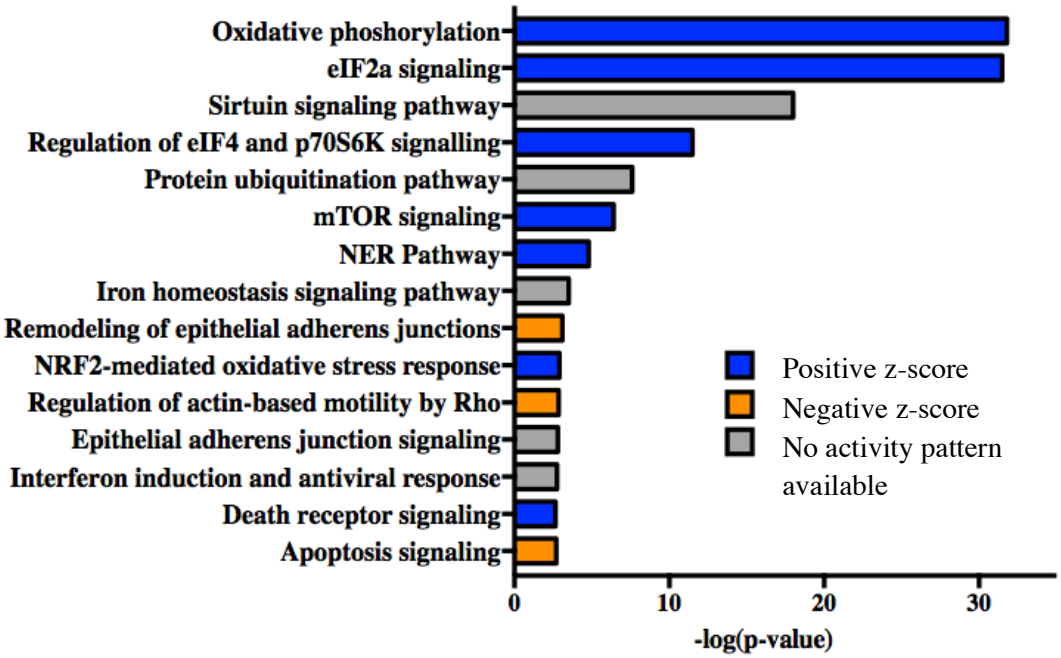


Figure 5.6: Top 15 biological pathways identified as being significantly altered between early Lyme disease patients and normal/healthy controls, as predicted by Ingenuity Pathway Analysis of DGE data. Calculated p-value (-log) is plotted along the x-axis for each pathway as calculated by IPA. Pathways with a positive z-score (decrease in activity in Lyme disease patients) are shown in blue, with negative z-score (increase in activity in Lyme disease patients) shown in orange. Where no activity pattern is available or can be calculated, grey bars are shown.

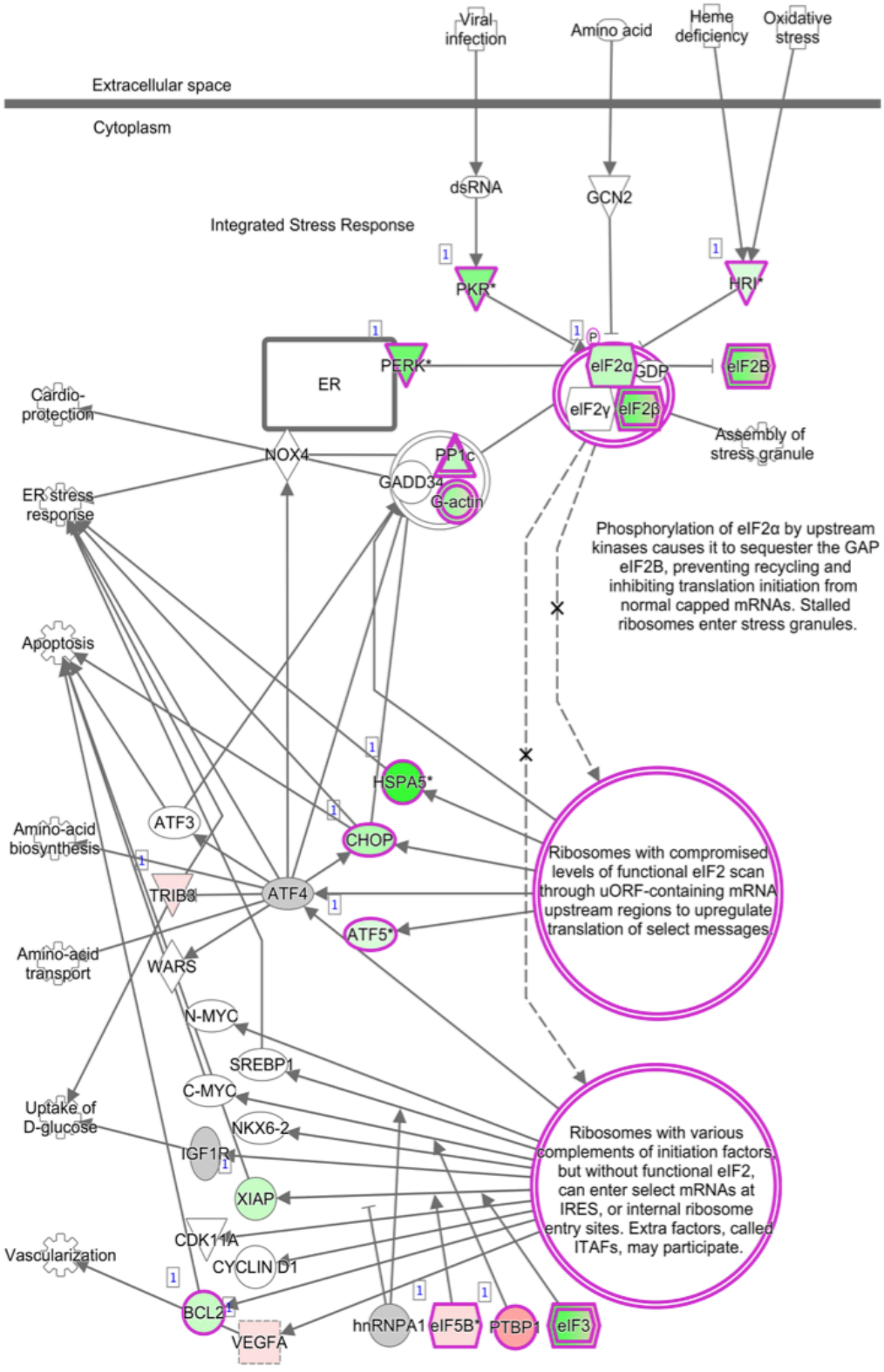


Figure 5.7: Ingenuity pathway analysis (IPA) network diagram of eIF2 illustrating annotated interactions between genes differentially expressed in Lyme disease patients against healthy controls. Molecules/genes identified as up regulated are shown in red, while those in green represent down regulated.

5.3.5. Ingenuity pathway analysis: Early vs. late Lyme disease

In order to identify pathways that may be increased in early Lyme disease, but decrease over the course of disease to convalescence, the DGE data generated in section 5.3.3 for "early" vs. "late" Lyme disease were input into IPA. No healthy controls were included in this analysis, and pathway changes represent only those between early and late disease, in this sample set. **Figure 5.8** shows 10 pathways identified by IPA between these groups. Of particular note is the significant increase seen in eIF2a signalling, suggesting that over the course of disease, this pathway reverts back to that seen in healthy patients. Again, oxidative phosphorylation as seen to be influenced between groups; in this case no particular activity pattern was identified. This result further suggests some mechanism occurring during early Lyme disease that disrupts or influences mitochondrial function in the human host. IL-10 and TLR signalling were identified as being decreased in activity in late Lyme disease, suggesting the down regulation of innate immunity and inflammatory processes back to normal levels. A small decrease in activity of actin-based motility by Rho was identified in late Lyme disease.

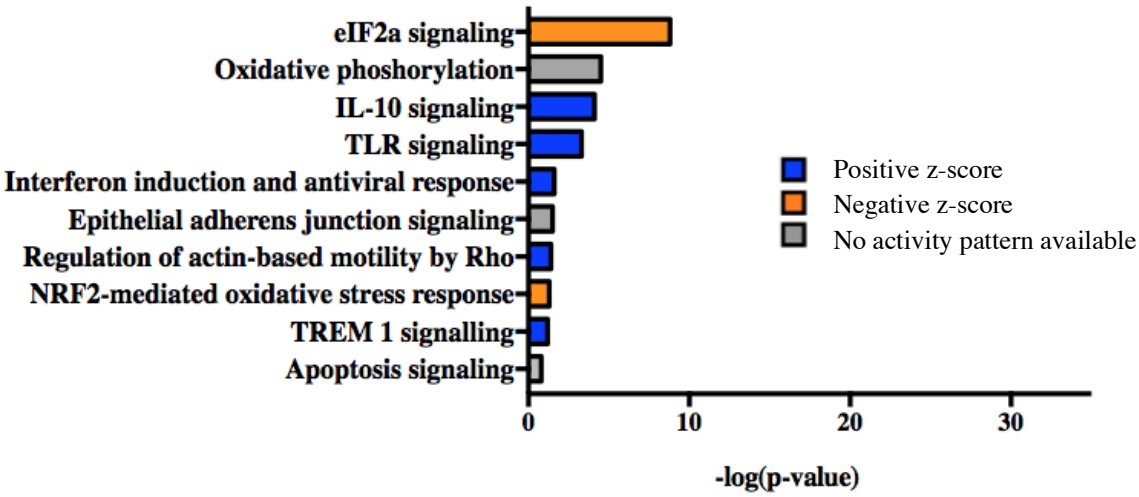


Figure 5.8: Top biological pathways detected by IPA software when comparing early Lyme disease to late Lyme disease, based on DGE data. Pathways with blue bars have a positive z-score (decrease in activity from early to late Lyme), while orange bars represent a negative z-score (increase in activity from early to late Lyme).

5.3.6. Cell typing analysis: ImmQuant

A final analysis based on the sequencing data for Lyme disease patients was conducted using omicX ImmQuant software (Materials and Methods 2.7.5). ImmQuant again uses DGE data to identify a list of specific genes that are associated with the presence of certain types of host immune cell. The software produces an estimation of the presence of cells of the presence together with a significance value for each. **Figure 5.9** shows the immune cells identified by ImmQuant as being significantly increased or decreased in early Lyme disease when compared to normal/healthy controls. Only cells with a significance of $p<0.05$ are shown. Erythroid CD34- CD71+ cells were found to have to have a large increase in presence in blood from Lyme disease patients compared to very low levels in normal/healthy controls. Myeloid dendritic cells (mDCs, also known as conventional dendritic cells) were found to be increased in abundance in the Lyme disease group. Dendritic cells

are antigen-presenting cells that have major importance in the mammalian immune system, allowing progress from innate to adaptive immune systems. Conversely, plasmacytoid dendritic cells (pDCs) were estimated to be decreased in abundance in the Lyme disease group. pDCs resemble plasma cells but share some features of mDCs including secretion of interferon-alpha and presentation of TLRs. Altered balance between pDCs and mDCs has been examined previously in other diseases, but an underlying biological reason for their contrasting abundance in early Lyme disease is difficult to explain without further study. Early and proto B-cells were found to be in lower abundance in Lyme disease patients, while mature B-cells were increased as compared to the control group. This finding suggests a decrease in proto-B cells as they progress into mature-B cells during Lyme disease.

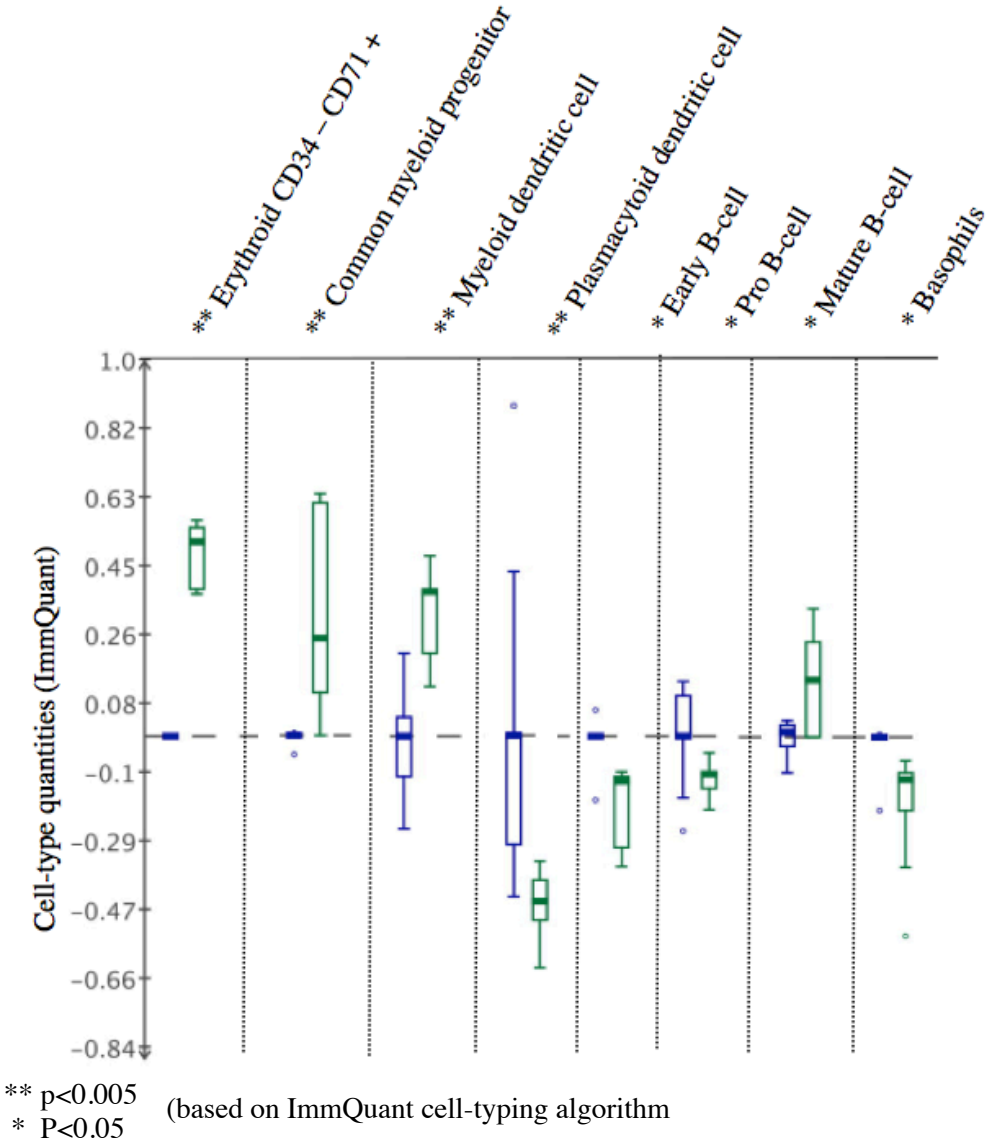


Figure 5.9: ImmQuant cell-typing data based on DGE analyses between early Lyme disease blood RNA to that of a normal/healthy control group. Bars in green represent the Lyme disease positive patient data, showing mean and SEM bars, with control data in blue. The Y-axis shows cell-type quantities as estimated by ImmQuant software (259).

5.3.7. RT-PCR analyses: eIF2A signalling in Lyme disease

To further examine expression of genes related to the eIF2A pathway, identified by IPA as significantly down regulated in early Lyme disease patients compared to normal/healthy controls, a RT-PCR based study was utilised. Primers were designed based on the mRNA sequence of several genes of interest. Sequence and design protocol for these primers is given in Materials and Methods 2.7.6. A total of 8 genes from the eIF2a signalling pathway were tested for in RNA extracts from early Lyme disease patients and compared to a set of 5 blood RNA extracts from normal/healthy controls. A quantitative PCR method was used (Materials and Methods 2.7.6). Full details and PCR cycling temperatures are provided in Materials and Methods. Cycle threshold (CT) was recorded per sample, and delta-delta-CT was calculated using the normal/healthy control group samples by means of a set of housekeeping genes. Of the 8 genes, 5 were successfully detected using RT-PCR in both Lyme disease patients and normal/healthy controls (**figure 5.10**). The genes AKT, eIF2a, eIF4g, XIAP and eIF5 were all shown to be increased or decreased in expression as previously detected using DGE.

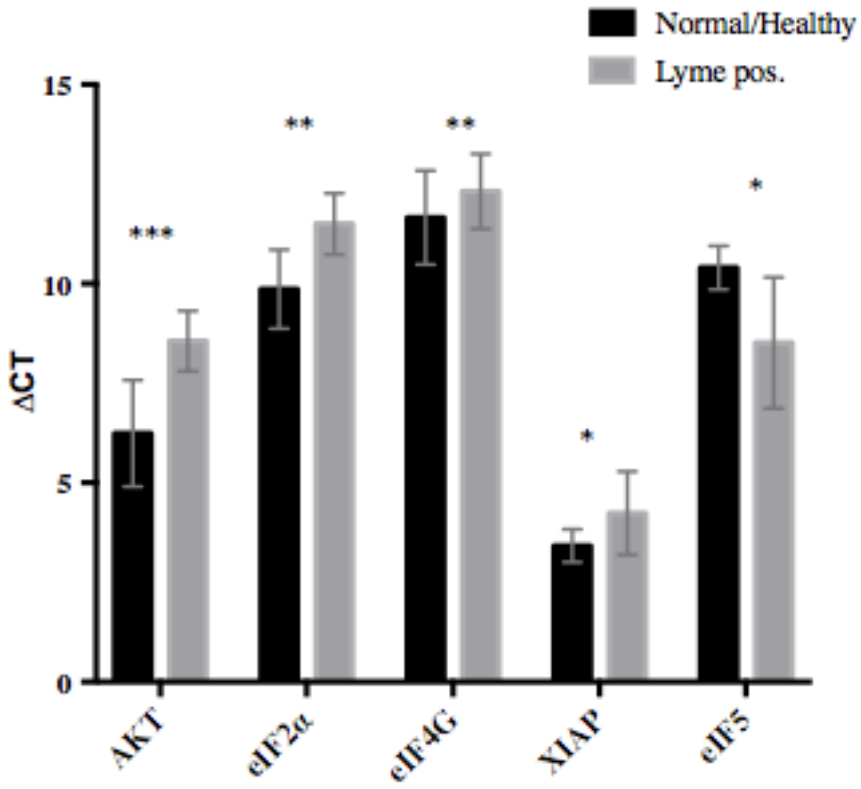


Figure 5.10: Graph showing RT-PCR data for 5 genes associated with eIF2a signalling in Lyme disease patients, compared with normal/healthy controls. Error bars represent standard error of the mean. *p<0.05, **p<0.005 ***p<0.0005

5.4. DISCUSSION

In this study, RNA was extracted from 14 whole blood samples from patients with Lyme disease, sequenced and compared to previously published reports of control groups comprising Ebola patients and normal/healthy controls. DGE analyses revealed large and consistent differences in blood transcriptome between early Lyme disease patients and healthy controls. These data were processed by a machine-learning algorithm to identify a panel of 9 genes that were together discriminatory of Lyme disease positive patients from controls in this dataset. Ingenuity Pathway Analysis (IPA) of the Lyme disease vs. normal/healthy controls DGE data identified a

number of pathways as being differentially expressed between groups. eIF2a signalling was found to be significantly decreased in Lyme disease patients by IPA. This supports previous findings of decreased activity of eIF2a signalling in Lyme disease (258). In eukaryotic cells, eIF2 (Eukaryotic Initiation Factor 2) is involved in translation initiation. eIF2 signalling has been shown to be induced following infection in response to endoplasmic reticulum stress, heavy metal toxicity and pathogen detection (260). Conversely, eIF2(a) signalling was found to be significantly decreased in Lyme disease patients by IPA in this study. The bacterial virulence factor YopJ has previously been shown to limit eIF2 signalling in mammalian cells, and was found to be of particular importance to several intracellular bacterial pathogens including *Listeria monocytogenes* and *Chlamydia trachomatis* (261). In contrast, the Lyme disease pathogen *Borrelia burgdorferi* s.l. does not express any known homologue of YopJ virulence factor; however, it can be hypothesised that the pathogen gains an evolutionary advantage from down regulation of eIF2 signalling due to decreased expression of proteins including those of the innate and adaptive response, and that some mechanism of bacterial influence over host signalling may exist. IPA also identified oxidative phosphorylation and several other pathways related to mitochondrial dysfunction as being influenced or disrupted in Lyme disease patients. Disruption of host phosphorylation has previously been described as a mechanism of pathogenesis, again by the intracellular pathogen *L. monocytogenes* (262). The immune effects of oxidative phosphorylation pathway are not clear, although previous studies have shown influence of neutrophils and T cells, reducing reactive oxygen species (ROS) production (263). This result again provides an example of the presumed extracellular pathogen behaving in a way associated with intracellular bacteria. No sound scientific evidence to date has shown *Borrelia* to act

as a intracellular pathogen during its life cycle, although this has been proposed (264, 265). eIF2(a) signalling was seen to increase back towards normal/healthy levels of expression in late Lyme disease patient samples. Actin rearrangement pathways were also identified as being altered in Lyme disease patients. Actin, tubulin and several other cell structural related proteins were also identified as being differentially abundant in sera of Lyme disease patients compared to controls by mass spectrometry (Chapters 3 and 4). As host defence mechanisms involve maintenance of numerous barriers to pathogens, including at the skin level and in connective tissues (266), disruption of cell structural rearrangement may confer a biological advantage to the infectious agent. Several examples also exist of intracellular bacterial pathogens effectively hijacking host cellular rearrangement mechanisms to physically transport themselves (267) or to promote internalisation into cells (268). As *Borrelia* has a well-described flagellar-based system for movement, the advantage to the pathogen of influencing host cytoskeletal rearrangement is unclear.

All analyses in this chapter were based on a set of 14 blood samples taken from 8 patients at different times after develop symptoms of Lyme disease. Due to time constraints, samples available during the early stages of the collaborative study with the colleagues in the Czech Republic were used. The main limitation of this study was a lack of consistent number of samples across time points, with some patients having undergone sampling at multiple time points, whilst others were sampled at only one time point with no RNA data available at time of presentation. No information about the age of the normal/healthy donors was available and these were matched based only on number of male/female samples included. Ebola acute and convalescent samples were included as an infectious disease control. The suitability of these as

controls may be questionable in this study due to the viral and severe nature of this infection, however they did provide some control against general immune response in the host. For future studies, a larger sample set could be utilised. Gene panels identified for early Lyme could be tested using well-defined Lyme disease samples, including from those that tested seronegative for Lyme using conventional diagnostic testing. A larger control group of samples, including appropriate infectious disease controls such as blood from syphilis patients would also be advantageous. Despite these limitations, the study was able to find significant and consistent changes at the transcriptome level in Lyme disease patients, and showed that these changes generally begin to revert back to normal/healthy levels late in the course of infection.

4.5.1. Conclusions

DGE analyses revealed large and consistent differences in blood transcriptome between early Lyme disease patients and healthy controls. A machine-learning algorithm to identify a panel of 9 genes that were together discriminatory of Lyme disease positive patients from controls. Ingenuity Pathway Analysis (IPA) of the Lyme disease vs. normal/healthy controls DGE data identified a number of differentially expressed pathways. Of note, eIF2a signalling was found to be significantly decreased in Lyme disease patients by IPA. Several findings of this study corroborate an earlier DGE study of Lyme disease, demonstrating the reproducibility of findings. Future host-gene expression analyses using well-characterised Lyme disease patient cohorts will help further our understanding of the complicated, multifaceted nature of the disease.

CHAPTER 6: CONCLUSIONS

The majority of Lyme disease diagnosis currently relies on the detection of host antibodies to *Borrelia burgdorferi* s.l. antigens in patient serum. Early in disease, when the humoral immune response is developing, serological tests have poor negative predictive value because borrelia-specific antibodies may be below detectable levels. *Borrelia*-specific antibodies may also persist for several years; therefore, in patients with persistent symptoms or suspected re-infection, positive serology does not necessarily indicate on-going infection. The aims of this thesis were to use quantitative proteomics and RNA sequencing to analyse the blood of Lyme disease patients, to make comparisons to normal healthy donors and related-disease control groups including those with syphilis and leptospirosis, and to identify potential biomarkers of early Lyme disease.

6.1. Proteomics-based investigation of Lyme disease and other spirochaetal infections

A proteomic approach using label-free mass spectrometry revealed very few differences in serum proteome between Lyme disease seropositive and seronegative groups. Differences observed between Lyme disease seropositive sera and normal healthy control sera showed that a range of innate and adaptive immune responses were occurring in the Lyme disease patients but, when compared to Seronegative, few of these changes were specific to the seropositive group. Possible reasons for this finding include that, because *B. burgdorferi* s.l. infection is marked by low bacteraemia that is initially localised to the site of tick bite with a relatively brief dissemination period, beyond triggering common host responses to bacterial infection,

the influence of Lyme disease on the serum proteome is modest in early disease. The findings are clouded by the fact that there is no real gold standard diagnostic test to confirm or rule out Lyme disease, particularly in the earliest stages, and that a patient that tests seronegative for Lyme disease may later seroconvert and test positive. Patients with documented retesting and seroconversion at a later date were not included in analyses, but in the majority of cases it was not possible to confirm whether the patient was subsequently diagnosed and treated for Lyme disease or a different ailment. A future study that attempted to identify protein signatures in the serum of early Lyme disease patients would ideally be a longitudinal study that involved retesting of all subjects at several time points.

To confirm that the quality of the residual diagnostic serum samples used in mass spectrometry run 1 was not responsible for the lack of difference observed, a second mass spectrometry run using samples taken at the time of presentation and immediately frozen was undertaken. The results of mass spectrometry run 2 again showed very few differences in protein abundance between Lyme disease seropositive and seronegative groups, and it was concluded that, at the resolution of mass spectrometry used in this study, the proteomes of individuals that test seropositive or seronegative for Lyme disease are highly similar. Of the few proteins found to be at significantly different abundance between Lyme disease seropositive and seronegative groups during mass spectrometry run 1, the protein Lipocalin 2 was of particular interest. The increased abundance of LCN2 in Lyme disease seropositive sera over seronegative was demonstrated by ELISA in additional samples (chapter 4). LCN2 is involved in a range of biological processes and is a well-described iron scavenger that is capable of removing iron from the site of bacterial infection so that the pathogens

are unable to meet their basic iron requirements for growth. Somewhat contradictory to this finding is the fact that *Borrelia* remain the only known organisms that can exist without iron (269). It is possible that LCN2 is induced as part of the innate response to borrelial infection despite the fact that it has little or no effect on the growth of *Borrelia*. The increased abundance of LCN2 in Lyme disease seropositive sera over seronegative was demonstrated by ELISA in additional samples; however, it should be noted that, as a protein often expressed during the immune response to infection, LCN2 was also elevated in leptospirosis and syphilis sera and was not specific to Lyme disease. As a proof-of-concept for the proteomic approach, a range of proteins were identified during mass spectrometry as differentially abundant in the sera of leptospirosis patients. The proteins serum amyloid A1 and carbonic anhydrase III were shown to be increased in leptospirosis positive sera by mass spectrometry and were both shown to be elevated above all other groups by ELISA in additional leptospirosis samples. The presence of elevated levels of the muscle-specific CAH3 in leptospirosis positive sera was hypothesised to be due to muscle damage during systemic infection and release of proteins into the bloodstream. Elevated CAH3 serum level has not previously been reported in leptospirosis.

6.2. Transcriptomic-based investigation of Lyme disease

A RNA-seq based study of the blood transcriptome of Lyme disease patients gave an insight into the host response to *Borrelia* infection. Studying a set of 14 early Lyme disease blood samples from Czech patients with an EM rash, clinically diagnosed Lyme disease and supporting laboratory serology, it was found that the eIF2 signalling was downregulated in Lyme disease patients. This is consistent with a finding that has previously been described in a transcriptomics study of Lyme disease in US cases,

suggesting this may be a shared response across infection with different borrelia genospecies (258). To investigate which responses may be limited to early disease, a comparison of early to late Lyme disease samples found that eIF2 signalling generally increased back towards levels found in normal healthy controls and therefore may be a potential indicator of successful treatment. RT-PCR analysis of 5 genes associated with eIF2 signalling were shown to be decreased in Lyme disease positive blood samples when compared to normal healthy controls. Ingenuity pathway analysis of the transcriptome data revealed a range of biological processes occurring in Lyme disease patients, from those associated with host immune responses to bacterial infection, to more unusual findings such as alterations in oxidative stress pathways and cell cytoskeletal rearrangement. The wealth of information provided by RNA-seq allows gene panels to be created that are highly specific to the exposure group and therefore the disease in question; however this process is sensitive to bias and could potentially reveal a gene panel that is diagnostic only of the input data. Any conclusions on the diagnostic potential of a gene panel would require extensive confirmatory testing in additional samples. A longitudinal study with regular sampling from well-defined Lyme disease patients could further identify genes and pathways that differentially expressed in early Lyme disease and revert back to normal levels following successful antibiotic treatment.

6.3. Comparison of proteomic and transcriptomic findings

Generally, proteomic findings were quite consistent with later transcriptomic data, with several of the proteins identified as being differentially abundant in Lyme also found to be differentially expressed; however, for some proteins, the link between expression and protein abundance was not as clear. RNA-seq involved sequencing of

mRNA from PBMCs and while proteins produced in these cells may be released into the serum, the serum also contains proteins produced in other areas of the body. It is also an overly simplistic view that all gene expression, as quantified by presence of specific mRNA, will result in a matching increase or decrease in protein abundance. A number of technical and biological factors are likely to contribute to the modest correlation seen between the transcriptome and proteome. At the transcript level, the effects of translational efficiency, RNA degradation, complex assembly, cell transport and alternative splicing all contribute to the tenuousness seen between the abundance of a particular mRNA transcript and translation into a protein. Further study into the correlation between the proteomic and transcriptomic findings presented here would be of interest and may help to elucidate discrepancies seen between data sets.

6.4. Translation of proteomic and transcriptomic findings

In chapters 3 and 4, the protein LCN2 was shown to be elevated in patients that were seropositive for Lyme disease over those that tested negative, both by mass spectrometry and subsequent WBs and ELISAs. As LCN2 was seen to be elevated in other infectious disease, it is likely that any diagnostic potential of LCN2 would be as part of a multifaceted test. Several other practical considerations also need to be made when assessing the suitability of a protein biomarker. Diagnostic samples taken at the point of primary care often need to be transported to other sites for centralised diagnosis, as is the case with Lyme disease. Any biomarker proteins would need to be quantifiable during diagnostic tests and therefore, labile proteins that are likely to degrade in conditions above freezing would not be suitable.

In chapter 5, differential gene expression data and a PLDA machine learning method identified a panel of 11 genes was identified as being discriminatory of early Lyme disease patient blood. Whilst the marker panel showed very high levels of sensitivity and specificity when tested in additional samples, the study was limited by a lack of available samples to serve as appropriate infectious controls. Theoretically, RT-PCR could be used to measure the expression of a specific gene panel but any transcriptomic based diagnostic of this type would also require whole blood to be collected and RNA stabilised, a process that isn't compatible with conventional diagnostic pathways and workflows; however, the transcriptomic data may be useful in identifying gene markers that could be verified at the protein level.

Throughout these studies emphasis was placed on early Lyme disease samples to try and identify signatures of the host response to active *Borrelia* infection. Subsequent studies to assess the potential of these markers would require testing in a large set of well characterised longitudinal patient samples to confirm whether they decline with treatment or in cases of spontaneous recovery and whether they are again elevated in cases of reinfection. This would be of particular interest in groups that are at high risk of Lyme including those with occupational exposures to ticks. Finally, the level of biomarkers associated with active *Borrelia* infection could be studied in patients with symptoms of post-treatment Lyme disease syndrome or apparent chronic Lyme disease, to investigate whether they remain elevated from early disease.

6.5. Final conclusions

In the world of human bacterial infections, Lyme disease is quite unusual. *Borrelia b*. s.l. begins preparing for invasion as soon as mammalian blood begins to flow into the midgut of an infected *Ixodes* tick whilst taking a blood-meal. Gene expression is altered as the spirochaete migrates through the tick vector, corkscrewing by use of intricate bundles of internal flagella, to the tick salivary glands. From here, it propels itself to the dermal skin layer of the human host, and colonisation can begin. The bacteria multiplies at the site of tick-bite and a characteristic erythema migrans rash may form; an unusual bullseye rash caused by the complex interplay between the multiplying invader and the inflammatory reaction of the host. From here, the timeline splits. A GP that recognises an EM rash on a patient presenting with a flu-like illness, who had been hiking in the Trossachs a week before, may immediately prescribe three weeks of doxycycline. If they are unsure, they can send a serum sample for diagnostic testing for Lyme disease. Serology results are taken into consideration, together with the clinical details of the case and, if deemed Lyme disease positive, antibiotics can be prescribed. Evidence shows that if diagnosed and treated at this early stage, bacterial-clearance and patient recovery rates are very high. In some cases, however, those infected with *Borrelia b*. s.l. may initially be asymptomatic, or show symptoms atypical of Lyme disease. Unrecognised and untreated, the infection can spread from the skin to the bloodstream, travel to various sites within the host, potentially causing late stage manifestations including Lyme arthritis, carditis and neuroborreliosis.

While the value of a diagnostic test capable of detecting Lyme disease at its earliest stages is clear, the nature of serological testing means that results are dependent on the humoral response of each individual being tested. Two-tiered testing algorithms

remain the most valuable tool for routine Lyme disease diagnostics and perform well in the majority of referrals. The protocols are being continually assessed and improved. However, issues remain around the efficacy of the test in the earliest stages of Lyme disease. Future proteomic, transcriptomic and metabolomic research may reveal further markers that will improve early diagnostic testing.

Late stage Lyme disease is often a complicated, multi-organ disease that is known to cause a wide range of symptoms, many of which are common to other diseases. As arguments continue over attributing Lyme disease (or Chronic Lyme disease) to a diverse range of cases, it should be remembered that there remains a group of individuals that feel that they have been let down; suffering from long-term maladies and unable to find a definitive diagnosis. Future understanding of these complicated cases lies in expanding our knowledge of both Lyme disease and other diseases, through scientific research, outreach and collaboration. “Never assume the obvious is true” (270).

REFERENCES

1. Gupta RS, Mahmood S, Adeolu M. *A phylogenomic and molecular signature based approach for characterization of the phylum Spirochaetes and its major clades: proposal for a taxonomic revision of the phylum*. Fr in microbiol. 2013;4:217.
2. Paster BJ, Dewhirst FE, Weisburg WG, Tordoff LA, Fraser GJ, Hespell RB, et al. *Phylogenetic analysis of the spirochetes.* J Bacteriol. 1991;173(19):6101-9.
3. Margos G, Gofton A, Wibberg D, Dangel A, Marosevic D, Loh SM, et al. *The genus Borrelia reloaded.* PLOS One. 2018;13(12):e0208432.
4. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. *The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks.* Nucleic Acids Res. 2014;42(Database issue):D643-D8.
5. Margos G, Lane RS, Fedorova N, Koloczek J, Piesman J, Hojgaard A, et al. *Borrelia bissettiae sp. nov. and Borrelia californiensis sp. nov. prevail in diverse enzootic transmission cycles.* Int J Syst Evol Microbiol. 2016;66(3):1447-52.
6. Eisen RJ, Eisen L. *The Blacklegged Tick, Ixodes scapularis: An Increasing Public Health Concern*. Trends Parisitol. 2018;34(4):295-309.
7. Steere AC, Malawista SE, Snydman DR, Shope RE, Andiman WA, Ross MR, et al. *Lyme arthritis: an epidemic of oligoarticular arthritis in children and adults in three connecticut communities.* Arthritis and rheumatism. 1977;20(1):7-17.
8. Burgdorfer W, Barbour AG, Hayes SF, Benach JL, Grunwaldt E, Davis JP. *Lyme disease-a tick-borne spirochetosis?* Science. 1982;216(4552):1317-9.
9. Alekseev AN, Dubinina HV, Van De Pol I, Schouls LM. *Identification of Ehrlichia spp. and Borrelia burgdorferi in Ixodes ticks in the Baltic regions of Russia.* J Clin Microbiol. 2001;39(6):2237-42.
10. Hansford KM, Fonville M, Gillingham EL, Coipan EC, Pietzsch ME, Krawczyk AI, et al. *Ticks and Borrelia in urban and peri-urban green space habitats in a city in southern England.* Ticks and tick-borne diseases. 2017;8(3):353-61.
11. Talleklint L, Jaenson TG. *Increasing geographical distribution and density of Ixodes ricinus (Acari: Ixodidae) in central and northern Sweden.* J. Med. Entomol. 1998;35(4):521-6.
12. Brownstein JS, Holford TR, Fish D. *Effect of Climate Change on Lyme Disease Risk in North America.* Ecohealth. 2005;2(1):38-46.
13. Rosenberg R, Lindsey NP, Fischer M, Gregory CJ, Hinckley AF, Mead PS, et al. Vital Signs: *Trends in Reported Vectorborne Disease Cases - United States and Territories, 2004-2016*. MMWR Morbidity and mortality weekly report. 2018;67(17):496-501.
14. Sykes RA, Makiello P. *An estimate of Lyme borreliosis incidence in Western Europedagger.* J. Pub. health. 2017;39(1):74-81.
15. Public Health England. *Lyme disease epidemiology and surveillance.* 2018. Available from https://www.gov.uk/government/publications/lyme-borreliosis-epidemiology/lyme-borreliosis-epidemiology-and-surveillance [accessed: 09/2019]
16. Veterinary Record. *Lyme disease cases on the rise in the UK*. Vet. Rec. 2018;183(2):44.
17. Tulloch JSP, Semper AE, Brooks TJG, Russell K, Halsby KD, Christley RM, et al. *The demographics and geographic distribution of laboratory-confirmed Lyme*

disease cases in England and Wales (2013–2016): an ecological study. BMJ Open. 2019;9(7):e028064.
18. Panetta JL, Šíma R, Calvani NED, Hajdušek O, Chandra S, Panuccio J, et al. *Reptile-associated Borrelia species in the goanna tick (Bothriocroton undatum) from Sydney, Australia.* Parasit Vectors. 2017;10(1):616.
19. Dworkin MS, Schwan TG, Anderson DE, Jr., Borchardt SM. *Tick-borne relapsing fever.* Infect Dis Clin North Am. 2008;22(3):449-viii.
20. Rawlings JA. *An overview of tick-borne relapsing fever with emphasis on outbreaks in Texas.* Texas medicine. 1995;91(5):56-9.
21. Trevejo RT, Schriefer ME, Gage KL, Safranek TJ, Orloski KA, Pape WJ, et al. *An interstate outbreak of tick-borne relapsing fever among vacationers at a Rocky Mountain cabin.* Am J Trop Med Hyg. 1998;58(6):743-7.
22. Schwan TG, Policastro PF, Miller Z, Thompson RL, Damrow T, Keirans JE. *Tick-borne relapsing fever caused by Borrelia hermsii, Montana.* Emerg Infect Dis. 2003;9(9):1151-4.
23. Thompson RS, Burgdorfer W, Russell R, Francis BJ. *Outbreak of tick-borne relapsing fever in Spokane County, Washington. Jama.* 1969;210(6):1045-50.
24. Boyer KM, Munford RS, Maupin GO, Pattison CP, Fox MD, Barnes AM, et al. *Tick-borne relapsing fever: an interstate outbreak originating at Grand Canyon National Park.* Am. J. Epidemol.. 1977;105(5):469-79.
25. Antinori S, Mediannikov O, Corbellino M, Grande R, Parravicini C, Bestetti G, et al. *Louse-Borne Relapsing Fever (Borrelia recurrentis) in a Somali Refugee Arriving in Italy: A Re-emerging Infection in Europe?* PLOS One. 2016;10(5):e0004522-e.
26. Seilmaier M, Guggemos W, Wieser A, Fingerle V, Balzer L, Fenzl T, et al. *[Louse-borne-relapsing-fever in refugees from the Horn of Africa; a case series of 25 patients]*. Deutsche medizinische Wochenschrift (1946). 2016;141(14):e133-42.
27. Ghazaei C. Pathogenic Leptospira: *Advances in understanding the molecular pathogenesis and virulence.* Open Vet J. 2018;8(1):13-24.
28. Casanovas-Massana A, Pedra GG, Wunder EA, Jr., Diggle PJ, Begon M, Ko AI. *Quantification of Leptospira interrogans Survival in Soil and Water Microcosms. Applied and environmental microbiology.* PLOS One. 2018;84(13).
29. Kobayashi Y. *Discovery of the causative organism of Weil's disease: historical view.* Journal of infection and chemotherapy : official journal of the Japan Society of Chemotherapy. 2001;7(1):10-5.
30. Gossman W, Stobart Gallagher MA, Dunn N. *Leptospirosis (Weil Disease).* StatPearls. Treasure Island FL: StatPearls Publishing LLC.; 2019.
31. Levett PN. *Leptospirosis*. Clin Microbiol Rev. 2001;14(2):296-326.
32. Abela-Ridder B, Sikkema R, Hartskeerl RA. E*stimating the burden of human leptospirosis. International journal of antimicrobial agents*. 2010;36 Suppl 1:S5-7.
33. Costa F, Hagan JE, Calcagno J, Kane M, Torgerson P, Martinez-Silveira MS, et al. *Global Morbidity and Mortality of Leptospirosis: A Systematic Review.* PLoS. 2015;9(9):e0003898.
34. Forbes AE, Zochowski WJ, Dubrey SW, Sivaprakasam V. *Leptospirosis and Weil's disease in the UK. QJM* : monthly journal of the Association of Physicians. 2012;105(12):1151-62.
35. Mitjà O, Šmajs D, Bassat Q. *Advances in the diagnosis of endemic treponematoses: yaws, bejel, and pinta.* PLoS. 2013;7(10):e2283-e.

36. Rothschild BM, Calderon FL, Coppa A, Rothschild C. *First European Exposure to Syphilis: The Dominican Republic at the Time of Columbian Contact.* Clin. Infec. 2000;31(4):936-41.
37. Kojima N, Klausner JD. *An Update on the Global Epidemiology of Syphilis.* Curr Epidemiol Rep. 2018;5(1):24-38.
38. Simms I, Tookey PA, Goh BT, Lyall H, Evans B, Townsend CL, et al. *The incidence of congenital syphilis in the United Kingdom: February 2010 to January 2015.* BJOG. 2017;124(1):72-7.
39. Iacobucci G. *Syphilis and gonorrhoea cases rose by a fifth in England last year.* BMJ. 2018;361:k2502.
40. Public Health England. *Sexually transmitted infections and screening for chlamydia in England, 2018 2019.* Avaiable from https://www.gov.uk/government/statistics/sexually-transmitted-infections-stis-annual-data-tables [Accessed: 12/2019]
41. Motaleb MA, Corum L, Bono JL, Elias AF, Rosa P, Samuels DS, et al. *Borrelia burgdorferi periplasmic flagella have both skeletal and motility functions. Proceedings of the National Academy of Sciences of the United States of America.* 2000;97(20):10899-904.
42. Kumar B, Miller K, Charon NW, Legleiter J. *Periplasmic flagella in Borrelia burgdoferi function to maintain cellular integrity upon external stress.* PLOS ONE. 2017;12(9):e0184648.
43. Ge Y, Li C, Corum L, Slaughter CA, Charon NW. *Structure and expression of the FlaA periplasmic flagellar protein of Borrelia burgdorferi.* J Bacteriol. 1998;180(9):2418-25.
44. Yang XF, Pal U, Alani SM, Fikrig E, Norgard MV. *Essential Role for OspA/B in the Life Cycle of the Lyme Disease Spirochete*. J. Exp. Med.. 2004;199(5):641.
45. Tilly K, Krum JG, Bestor A, Jewett MW, Grimm D, Bueschel D, et al. *Borrelia burgdorferi OspC protein required exclusively in a crucial early stage of mammalian infection*. Infection and immunity. 2006;74(6):3554-64.
46. Casjens SR, Mongodin EF, Qiu W-G, Luft BJ, Schutzer SE, Gilcrease EB, et al. *Genome Stability of Lyme Disease Spirochetes: Comparative Genomics of Borrelia burgdorferi Plasmids.* PLOS ONE. 2012;7(3):e33280.
47. Beaurepaire C, Chaconas G. *Topology-dependent transcription in linear and circular plasmids of the segmented genome of Borrelia burgdorferi.* Mol. microbiol. 2007;63(2):443-53.
48. Casjens S, Palmer N, van Vugt R, Huang WM, Stevenson B, Rosa P, et al. *A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete Borrelia burgdorferi.* Mol. microbiol. 2000;35(3):490-516.
49. Saier MH, Jr., Paulsen IT. *Whole genome analyses of transporters in spirochetes: Borrelia burgdorferi and Treponema pallidum.* J Mol. microbiol. Biotech. 2000;2(4):393-9.
50. Brisson D, Drecktrah D, Eggers CH, Samuels DS. *Genetics of Borrelia burgdorferi.* Annu Rev Genet. 2012;46:515-36.
51. Posey JE, Gherardini FC. *Lack of a Role for Iron in the Lyme Disease Pathogen.* Science. 2000;288(5471):1651.
52. Xiao X, Yeoh BS, Vijay-Kumar M. *Lipocalin 2: An Emerging Player in Iron Homeostasis and Inflammation.* Annual review of nutrition. 2017;37:103-30.

53. Troxell B, Xu H, Yang XF. *Borrelia burgdorferi, a pathogen that lacks iron, encodes manganese-dependent superoxide dismutase essential for resistance to streptonigrin.* J Biol Chem. 2012;287(23):19284-93.
54. Lambert A, Picardeau M, Haake DA, Sermswan RW, Srikram A, Adler B, et al. *FlaA proteins in Leptospira interrogans are essential for motility and virulence but are not required for formation of the flagellum sheath*. Inf. and imm. 2012;80(6):2019-25.
55. Verma A, Hellwage J, Artiushin S, Zipfel PF, Kraiczy P, Timoney JF, et al. *LfhA, a novel factor H-binding protein of Leptospira interrogans.* Inf. and imm. 2006;74(5):2659-66.
56. Jorge S, Kremer FS, Oliveira NRd, Navarro GdOSV, Guimarães AM, Sanchez CD, et al. *Whole-genome sequencing of Leptospira interrogans from southern Brazil: genetic features of a highly virulent strain*. Mem Inst Oswaldo Cruz. 2018;113(2):80-6.
57. Picardeau M, Bulach DM, Bouchier C, Zuerner RL, Zidane N, Wilson PJ, et al. *Genome sequence of the saprophyte Leptospira biflexa provides insights into the evolution of Leptospira and the pathogenesis of leptospirosis.* PLOS One. 2008;3(2):e1607.
58. Blanco DR, Miller JN, Lovett MA. S*urface antigens of the syphilis spirochete and their potential as virulence determinants.* Emerging infectious diseases. 1997;3(1):11-20.
59. Fraser CM, Norris SJ, Weinstock GM, White O, Sutton GG, Dodson R, et al. *Complete genome sequence of Treponema pallidum, the syphilis spirochete.* Science. 1998;281(5375):375-88.
60. LaFond RE, Lukehart SA. *Biological Basis for Syphilis.* Clin. Microbiol. 2006;19(1):29.
61. Nascimento ALTO, Verjovski-Almeida S, Van Sluys MA, Monteiro-Vitorello CB, Camargo LEA, Digiampietri LA, et al. *Genome features of Leptospira interrogans serovar Copenhageni.* Braz J Med Biol Res. 2004;37(4):459-77.
62. Estrada-Pena A, Cutler S, Potkonjak A, Vassier-Tussaut M, Van Bortel W, Zeller H, et al. *An updated meta-analysis of the distribution and prevalence of Borrelia burgdorferi s.l. in ticks in Europe.* International journal of health geographics. 2018;17(1):41.
63. Nelson C, Banks S, Jeffries CL, Walker T, Logan JG. T*ick abundances in South London parks and the potential risk for Lyme borreliosis to the general public.* Medical and veterinary entomology. 2015;29(4):448-52.
64. Wikel SK. *Ticks and Tick-Borne Infections: Complex Ecology, Agents, and Host Interactions*. Vet Sci. 2018;5(2):60.
65. Pietzsch ME, Medlock JM, Jones L, Avenell D, Abbott J, Harding P, et al. *Distribution of Ixodes ricinus in the British Isles: investigation of historical records.* Med. entomol. 2005;19(3):306-14.
66. Scharlemann JPW, Johnson PJ, Smith AA, Macdonald DW, Randolph SE. *Trends in ixodid tick abundance and distribution in Great Britain.* Med. entomol. 2008;22(3):238-47.
67. Cull B, Pietzsch ME, Hansford KM, Gillingham EL, Medlock JM. *Surveillance of British ticks: An overview of species records, host associations, and new records of Ixodes ricinus distribution.* Ticks and tick-borne diseases. 2018;9(3):605-14.

68. Radolf JD, Caimano MJ, Stevenson B, Hu LT. *Of ticks, mice and men: understanding the dual-host lifestyle of Lyme disease spirochaetes*. Nat Rev Microbiol. 2012;10(2):87-99.
69. Nieto NC, Porter WT, Wachara JC, Lowrey TJ, Martin L, Motyka PJ, et al. *Using citizen science to describe the prevalence and distribution of tick bite and exposure to tick-borne diseases in the United States.* PLOS One 2018;13(7):e0199644.
70. Medlock JM, Hansford KM, Vaux AGC, Cull B, Gillingham E, Leach S. *Assessment of the Public Health Threats Posed by Vector-Borne Disease in the United Kingdom (UK).* International journal of environmental research and public health. 2018;15(10).
71. Eisen L. *Pathogen transmission in relation to duration of attachment by Ixodes scapularis ticks.* Ticks and tick-borne diseases. 2018;9(3):535-42.
72. Rumer L, Sheshukova O, Dautel H, Donoso Mantke O, Niedrig M. *Differentiation of medically important Euro-Asian tick species Ixodes ricinus, Ixodes persulcatus, Ixodes hexagonus, and Dermacentor reticulatus by polymerase chain reaction*. Vector borne and zoonotic diseases. 2011;11(7):899-905.
73. Clark RP, Hu LT. *Prevention of lyme disease and other tick-borne infections.* Infect Dis Clin North Am. 2008;22(3):381-vii.
74. Embers ME, Narasimhan S. *Vaccination against Lyme disease: past, present, and future.* Front. Cell. Infec. Biol. 2013;3:6-.
75. Aronowitz RA. *The rise and fall of the lyme disease vaccines: a cautionary tale for risk interventions in American medicine and public health*. Milbank Q. 2012;90(2):250-77.
76. Nigrovic LE, Thompson KM. *The Lyme vaccine: a cautionary tale*. Epidemiol Infect. 2007;135(1):1-8.
77. Haake DA, Levett PN. *Leptospirosis in humans*. Curr Top Microbiol Immunol. 2015;387:65-97.
78. Casanovas-Massana A, Pedra GG, Wunder EA, Diggle PJ, Begon M, Ko AI. *Leptospira interrogans; Survival in Soil and Water Microcosms*. App. enviro. Biol. 2018;84(13):e00507-18.
79. Wynwood SJ, Graham GC, Weier SL, Collet TA, McKay DB, Craig SB. *Leptospirosis from water sources*. Pathog Glob Health. 2014;108(7):334-8.
80. Blasdell KR, Morand S, Perera D, Firth C. *Association of rodent-borne Leptospira spp. with urban environments in Malaysian Borneo.* PLoS neglected tropical diseases. 2019;13(2):e0007141.
81. Owusu-Ofori AK, Parry CM, Bates I. *Transfusion-transmitted syphilis in teaching hospital, Ghana.* Emerging infectious diseases. 2011;17(11):2080-2.
82. Chambers RW, Foley HT, Schmidt PJ. *Transmission of syphilis by fresh blood components.* Transfusion. 1969;9(1):32-4.
83. Stoltey JE, Cohen SE. *Syphilis transmission: a review of the current evidence.* Sex Health. 2015;12(2):103-9.
84. Arnold SR, Ford-Jones EL. *Congenital syphilis: A guide to diagnosis and management.* Paediatr Child Health. 2000;5(8):463-9.
85. Serpa JA, Huynh GN, Nickell JB, Miao H. *HIV Pre-Exposure Prophylaxis and Increased Incidence of Sexually Transmitted Infections in the United States. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*. 2019.

86. Samuels DS. *Gene regulation in Borrelia burgdorferi.* Ann. Rev. microbiol. 2011;65:479-99.
87. Indest KJ, Howell JK, Jacobs MB, Scholl-Meeker D, Norris SJ, Philipp MT. *Analysis of Borrelia burgdorferi vlsE gene expression and recombination in the tick vector*. Infect Immun. 2001;69(11):7083-90.
88. Tilly K, Rosa PA, Stewart PE. *Biology of infection with Borrelia burgdorferi.* Infect Dis Clin North Am. 2008;22(2):217-v.
89. Johnson KO, Nelder MP, Russell C, Li Y, Badiani T, Sander B, et al. *Clinical manifestations of reported Lyme disease cases in Ontario, Canada*: 2005–2014. PLOS One. 2018;13(6):e0198509.
90. NICE. *NICE guideline: Lyme disease 2018.* Available at https://www.nice.org.uk/guidance/ng95 [Accessed: 01/2019]
91. Kowalski TJ, Tata S, Berth W, Mathiason MA, Agger WA. *Antibiotic Treatment Duration and Long-Term Outcomes of Patients with Early Lyme Disease from a Lyme Disease–Hyperendemic Area*. Cl. I. Diseases. 2010;50(4):512-20.
92. Garcia-Monco JC, Benach JL. T*he pathogenesis of Lyme disease. Rheumatic diseases clinics of North America.* 1989;15(4):711-26.
93. Eldin C, Raffetin A, Bouiller K, Hansmann Y, Roblot F, Raoult D, et al. *Review of European and American guidelines for the diagnosis of Lyme borreliosis.* Medecine et maladies infectieuses. 2019;49(2):121-32.
94. Salzman BE, Stonehouse A, Studdiford J. *Late Diagnosis of Early Disseminated Lyme Disease: Perplexing Symptoms in a Gardener.* The Journal of the American Board of Family Medicine. 2008;21(3):234.
95. Pachner AR, Steiner I. *Lyme neuroborreliosis: infection, immunity, and inflammation.* The Lancet Neurology. 2007;6(6):544-52.
96. Sigal LH. *Early disseminated Lyme disease: Cardiac manifestations*. The Am. J of Med. 1995;98(4):25S-9S.
97. Murray TS, Shapiro ED. *Lyme disease*. Clin Lab Med. 2010;30(1):311-28.
98. Bockenstedt LK, Wormser GP. *Review: unraveling Lyme disease.* Arthritis & rheumatology. 2014;66(9):2313-23.
99. Steere AC, Schoen RT, Taylor E. *The clinical evolution of Lyme arthritis. Annals of internal medicine.* 1987;107(5):725-31.
100. Nygard K, Brantsaeter AB, Mehl R. *Disseminated and chronic Lyme borreliosis in Norway, 1995 - 2004. Euro surveillance* : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2005;10(10):235-8.
101. Steere AC, Malawista SE, Hardin JA, Ruddy S, Askenase W, Andiman WA. *Erythema chronicum migrans and Lyme arthritis. The enlarging clinical spectrum*. Annals of internal medicine. 1977;86(6):685-98.
102. Oschmann P, Dorndorf W, Hornig C, Schafer C, Wellensiek HJ, Pflughaupt KW. *Stages and syndromes of neuroborreliosis*. Journal of neurology. 1998;245(5):262-72.
103. Dumic I, Vitorovic D, Spritzer S, Sviggum E, Patel J, Ramanan P. *Acute transverse myelitis - A rare clinical manifestation of Lyme neuroborreliosis*. IDCases. 2018;15:e00479-e.
104. Halperin JJ. *Diagnosis and management of Lyme neuroborreliosi*s. Expert review of anti-infective therapy. 2018;16(1):5-11.
105. Cairns V, Godwin J. P*ost-Lyme borreliosis syndrome: a meta-analysis of reported symptoms.* International journal of epidemiology. 2005;34(6):1340-5.

106. Vrijmoeth HD, Ursinus J, Harms MG, Zomer TP, Gauw SA, Tulen AD, et al. *Prevalence and determinants of persistent symptoms after treatment for Lyme borreliosis: study protocol for an observational, prospective cohort study (LymeProspect).* BMC infectious diseases. 2019;19(1):324.
107. (CDC) CfDCaP. *Post-Treatment Lyme Disease Syndrome*. 2019. Available at: https://www.cdc.gov/lyme/postlds/index.html [Accessed: 06/2019]
108. Weinstein A, Britchkov M. Lyme arthritis and post-Lyme disease syndrome. Current opinion in rheumatology. 2002;14(4):383-7.
109. Society ILaAD. *Controversy over Chronic Lyme Disease 2019* [Available from: https://www.ilads.org/research-literature/controversies-challenges/. [Accessed 12/2019]
110. Harley JB, Chen X, Pujato M, Miller D, Maddox A, Forney C, et al. *Transcription factors operate across disease loci, with EBNA2 implicated in autoimmunity.* Nat Genet. 2018;50(5):699-707.
111. Ma Y, Sturrock A, Weis JJ. *Intracellular localization of Borrelia burgdorferi within human endothelial cells.* Infection and immunity. 1991;59(2):671-8.
112. Cabello FC, Godfrey HP, Newman SA. *Hidden in plain sight: Borrelia burgdorferi and the extracellular matrix.* Trends in microbiology. 2007;15(8):350-4.
113. Brown EL, Wooten RM, Johnson BJ, Iozzo RV, Smith A, Dolan MC, et al. *Resistance to Lyme disease in decorin-deficient mice.* The Journal of clinical investigation. 2001;107(7):845-52.
114. Müller KE. *Damage of collagen and elastic fibres by borrelia burgdorferi - known and new clinical and histopathological aspects.* The open neurology journal. 2012;6:179-86.
115. Chong R, Swiss R, Briones G, Stone KL, Gulcicek EE, Agaisse H. *Regulatory mimicry in Listeria monocytogenes actin-based motility*. Cell Host Microbe. 2009;6(3):268-78.
116. Heinzen RA, Grieshaber SS, Van Kirk LS, Devin CJ. *Dynamics of actin-based movement by Rickettsia rickettsii in vero cells.* Infection and immunity. 1999;67(8):4201-7.
117. Lee JH, Park H, Park YH. *Molecular mechanisms of host cytoskeletal rearrangements by Shigella invasins*. Int J Mol Sci. 2014;15(10):18253-66.
118. Wu J, Weening EH, Faske JB, Höök M, Skare JT. *Invasion of Eukaryotic Cells.* Infection and Immunity. 2011;79(3):1338.
119. Schutze GE, Regan J, Bradley J. *Use doxycycline as first-line treatment for rickettsial diseases*. AAP News. 2010;31(7):14.
120. Holmes NE, Charles PGP. *Safety and Efficacy Review of Doxycycline.* Clinical Medicine Therapeutics. 2009;1:CMT.S2035.
121. Stricker RB, Delong AK, Green CL, Savely VR, Chamallas SN, Johnson L. *Benefit of intravenous antibiotic therapy in patients referred for treatment of neurologic Lyme disease.* International journal of general medicine. 2011;4:639-46.
122. Marks CM, Nawn JE, Caplow JA. *Antibiotic Treatment for Chronic Lyme Disease—Say No to the DRESS.* JAMA Internal Medicine. 2016;176(12):1745-6.
123. Berende A, ter Hofstede HJM, Vos FJ, van Middendorp H, Vogelaar ML, Tromp M, et al. *Randomized Trial of Longer-Term Therapy for Symptoms Attributed to Lyme Disease.* New England Journal of Medicine. 2016;374(13):1209-20.

124. Marzec NS, Nelson C, Waldron PR, Blackburn BG, Hosain S, Greenhow T, et al. *Serious Bacterial Infections Acquired During Treatment of Patients Given a Diagnosis of Chronic Lyme Disease - United States.* MMWR Morbidity and mortality weekly report. 2017;66(23):607-9.
125. Rudenko N, Golovchenko M, Kybicova K, Vancova M. *Metamorphoses of Lyme disease spirochetes: phenomenon of Borrelia persisters*. Parasit Vectors. 2019;12(1):237.
126. Christian P. *Revue critique des études essayant d'évaluer le traitement de la maladie de Lyme chronique.* La Presse Médicale. 2015;44.
127. Wormser GP, Dattwyler RJ, Shapiro ED, Halperin JJ, Steere AC, Klempner MS, et al. T*he clinical assessment, treatment, and prevention of lyme disease, human granulocytic anaplasmosis, and babesiosis: clinical practice guidelines by the Infectious Diseases Society of America*. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2006;43(9):1089-134.
128. Mygland A, Ljostad U, Fingerle V, Rupprecht T, Schmutzhard E, Steiner I, et al. *EFNS guidelines on the diagnosis and management of European Lyme neuroborreliosis.* Eur J Neurol. 2010;17(1):8-16, e1-4.
129. Flisiak R, Pancewicz S. *Diagnostyka i leczenie boreliozy z Lyme. Zalecenia Polskiego Towarzystwa Epidemiologicznego i Chorób Zakaźnych* [in Polish]. Przeglad epidemiologiczny. 2008;62(1):193-9.
130. Evison J, Aebi C, Francioli P, Peter O, Bassetti S, Gervaix A, et al. *Maladie de Lyme Partie I: épidémiologie et diagnostic* [in French]. Revue medicale suisse. 2006;2(60):919-24.
131. Dessau RB, Bangsborg JM, Jensen TP, Hansen K, Lebech AM, Andersen CO. *Laboratoriediagnose af infektion forårsaget af Borrelia burgdorferi* [in Danish]. Ugeskrift for laeger. 2006;168(34):2805-7.
132. Neurologie DGf. *Neuroborreliose* [in German] 2019 [Available from: https://www.dgn.org/leitlinien/3567-ll-030-071-2018-neuroborreliose.
133. Rayment C, O'Flynn N. *Diagnosis and management of patients with Lyme disease: NICE guideline*. The British journal of general practice : the journal of the Royal College of General Practitioners. 2018;68(676):546-7.
134. Halperin JJ, Shapiro ED, Logigian E, Belman AL, Dotevall L, Wormser GP, et al. *Practice parameter: treatment of nervous system Lyme disease (an evidence-based review):* report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology. 2007;69(1):91-102.
135. Stanek G, O'Connell S, Cimmino M, Aberer E, Kristoferitsch W, Granstrom M, et al. *European Union Concerted Action on Risk Assessment in Lyme Borreliosis: clinical case definitions for Lyme borreliosis*. Wiener klinische Wochenschrift. 1996;108(23):741-7.
136. Kimberlin DW. *Red Book: 2018-2021 report of the committee on infectious diseases.* American Academy of Pediatrics. 2020.
137. Marzec NS NC, Waldron PR, et al. . *Serious Bacterial Infections Acquired During Treatment of Patients Given a Diagnosis of Chronic Lyme Disease — United States*. CDC MMWR Morb Mortal Wkly Rep. 2017;Wkly Rep 2017;66:607–609.
138. Phraisuwan P, Whitney EAS, Tharmaphornpilas P, Guharat S, Thongkamsamut S, Aresagig S, et al. *Leptospirosis: skin wounds and control strategies, Thailand,* 1999. Emerging infectious diseases. 2002;8(12):1455-9.
139. Abdulkader RC, Silva MV. *The kidney in leptospirosis*. Pediatric nephrology (Berlin, Germany). 2008;23(12):2111-20.

140. Gonçalves-de-Albuquerque CF, Burth P, Silva AR, Younes-Ibrahim M, Castro-Faria-Neto HC, Castro-Faria MV. *Leptospira and inflammation.* Mediators of inflammation. 2012;2012:317950-.
141. Cagliero J, Villanueva SYAM, Matsui M. *Leptospirosis Pathophysiology: Into the Storm of Cytokines.* Fr. in microbiol. 2018;8:204-.
142. Gomes-Solecki M, Santecchia I, Werts C. *Animal Models of Leptospirosis: Of Mice and Hamsters.* Fr. in immunology. 2017;8:58-.
143. Ramachandran S, Rajapakse CN, Perera MV, Yoganathan M. *Patterns of acute renal failure in leptospirosis*. The Journal of tropical medicine and hygiene. 1976;79(7):158-60.
144. Chang MY, Cheng YC, Hsu SH, Ma TL, Chou LF, Hsu HH, et a*l. Leptospiral outer membrane protein LipL32 induces inflammation and kidney injury in zebrafish larvae.* Scientific reports. 2016;6:27838.
145. Levett PN. *Leptospirosis.* Clinical Microbiology Reviews. 2001;14(2):296.
146. De Brito T, Silva AMGd, Abreu PAE. *Pathology and pathogenesis of human leptospirosis: a commented review*. Rev Inst Med Trop Sao Paulo. 2018;60:e23-e.
147. Sperber SJ, Schleupner CJ. *Leptospirosis: a forgotten cause of aseptic meningitis and multisystem febrile illness*. South Med J. 1989;82(10):1285-8.
148. Peeling RW, Hook EW, 3rd. *The pathogenesis of syphilis: the Great Mimicker, revisited.* The Journal of pathology. 2006;208(2):224-32.
149. Kent ME, Romanelli F. *Reexamining syphilis: an update on epidemiology, clinical manifestations, and management.* The Annals of pharmacotherapy. 2008;42(2):226-36.
150. French P. *Syphilis.* BMJ. 2007;334(7585):143-7.
151. Pastuszczak M, Wojas-Pelc A. *Current standards for diagnosis and treatment of syphilis: selection of some practical issues, based on the European (IUSTI) and U.S. (CDC) guidelines*. Postepy Dermatol Alergol. 2013;30(4):203-10.
152. Agampodi SB, Dahanayaka NJ, Nöckler K, Mayer-Scholl A, Vinetz JM. *Redefining Gold Standard Testing for Diagnosing Leptospirosis: Further Evidence from a Well-Characterized, Flood-Related Outbreak in Sri Lanka.* Am J Trop Med Hyg. 2016;95(3):531-6.
153. Winslow WE, Merry DJ, Pirc ML, Devine PL. *Evaluation of a commercial enzyme-linked immunosorbent assay for detection of immunoglobulin M antibody in diagnosis of human leptospiral infection*. J Clin Microbiol. 1997;35(8):1938-42.
154. Cumberland P, Everard CO, Wheeler JG, Levett PN. *Persistence of anti-leptospiral IgM, IgG and agglutinating antibodies in patients presenting with acute febrile illness in Barbados 1979-1989.* European journal of epidemiology. 2001;17(7):601-8.
155. Woods K, Nic-Fhogartaigh C, Arnold C, Boutthasavong L, Phuklia W, Lim C, et al. *A comparison of two molecular methods for diagnosing leptospirosis from three different sample types in patients presenting with fever in Laos.* Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases. 2018;24(9):1017 e1- e7.
156. Musso D, La Scola B. *Laboratory diagnosis of leptospirosis: a challenge. Journal of microbiology, immunology, and infection* = Wei mian yu gan ran za zhi. 2013;46(4):245-52.
157. Turner LH. *Leptospirosis. 3. Maintenance, isolation and demonstration of leptospires*. Transactions of the Royal Society of Tropical Medicine and Hygiene. 1970;64(4):623-46.

158. Morshed MG, Singh AE. *Recent Trends in the Serologic Diagnosis of Syphilis.* Clinical and Vaccine Immunology. 2015;22(2):137.
159. Manavi K, Young H, McMillan A. T*he sensitivity of syphilis assays in detecting different stages of early syphilis.* International journal of STD & AIDS. 2006;17(11):768-71.
160. Kaur G, Kaur P. *Syphilis testing in blood donors: an update*. Blood Transfus. 2015;13(2):197-204.
161. Public Health England. *UK Standards for Microbiology Investigations 2016.* Available at: https://www.gov.uk/government/collections/standards-for-microbiology-investigations-smi [accessed 04/2019]

162. Garakani A, Mitton AG. *New-Onset Panic, Depression with Suicidal Thoughts, and Somatic Symptoms in a Patient with a History of Lyme Disease.* Case Reports in Psychiatry. 2015;2015:4.
163. Gendelman H, Schutzer SE, Angel TE, Liu T, Schepmoes AA, Clauss TR, et al. *Distinct Cerebrospinal Fluid Proteomes Differentiate Post-Treatment Lyme Disease from Chronic Fatigue Syndrome.* PLOS One. 2011;6(2):e17287.
164. Gautam A, Dixit S, Philipp MT, Singh SR, Morici LA, Kaushal D, et al. *Interleukin-10 Alters Effector Functions of Multiple Genes Induced by Borrelia burgdorferi in Macrophages To Regulate Lyme Disease Inflammation.* Infection and Immunity. 2011;79(12):4876-92.
165. Halperin J. *Nervous system Lyme disease, chronic Lyme disease, and none of the above.* Acta Neurologica Belgica. 2015:1-6.
166. Hassett AL, Radvanski DC, Buyske S, Savage SV, Gara M, Escobar JI, et al. *Role of psychiatric comorbidity in chronic Lyme disease*. Arthritis Rheum. 2008;59(12):1742-9.
167. Henry B, Roth D, Reilly R, MacDougall L, Mak S, Li M, et al. *How big is the Lyme problem? Using novel methods to estimate the true number of Lyme disease cases in British Columbia residents from 1997 to 2008.* Vector Borne Zoonotic Dis. 2011;11(7):863-8.
168. Hinckley AF, Connally NP, Meek JI, Johnson BJ, Kemperman MM, Feldman KA, et al. *Lyme Disease Testing by Large Commercial Laboratories in the United States. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2014;59(5):676-81.
169. Pomelova V, Korenberg E, Kuznetsova T, Osin N. *Utility of Borrelia burgdorferi sensu stricto C6 Peptide for Serologic Confirmation of Erythema-Free Ixodid Tick-Borne Borrelioses in Russia*. Biomed Res Int. 2018;2018:5291926-.
170. Miraglia CM. *A Review of the Centers for Disease Control and Prevention's Guidelines for the Clinical Laboratory Diagnosis of Lyme Disease.* J Chiropr Med. 2016;15(4):272-80.
171. Aguero-Rosenfeld ME, Wormser GP. *Lyme disease: diagnostic issues and controversies.* Expert Review of Molecular Diagnostics. 2015;15(1):1-4.
172. Aguero-Rosenfeld ME, Nowakowski J, McKenna DF, Carbonaro CA, Wormser GP. *Serodiagnosis in early Lyme disease.* J Clin Microbiol. 1993;31(12):3090-5.
173. Aguero-Rosenfeld ME, Nowakowski J, Bittker S, Cooper D, Nadelman RB, Wormser GP. *Evolution of the serologic response to Borrelia burgdorferi in treated patients with culture-confirmed erythema migrans.* J Clin Microbiol. 1996;34(1):1-9.

174. Engstrom SM, Shoop E, Johnson RC*. Immunoblot interpretation criteria for serodiagnosis of early Lyme disease.* J Clin Microbiol. 1995;33(2):419-27.
175. Evans R. MS, Ho-Yen D. *Audit of the laboratory diagnosis of Lyme disease in Scotland*. J Med Microbiol. (2005 Dec;54(Pt 12):1139-41.).
176. Cerar T, Strle F, Stupica D, Ruzic-Sabljic E, McHugh G, Steere AC, et al. *Differences in Genotype, Clinical Features, and Inflammatory Potential of Borrelia burgdorferi sensu stricto Strains from Europe and the United States.* Emerg Infect Dis. 2016;22(5):818-27.
177. Wormser GP, Liveris D, Hanincova K, Brisson D, Ludin S, Stracuzzi VJ, et al. *Effect of Borrelia burgdorferi genotype on the sensitivity of C6 and 2-tier testing in North American patients with culture-confirmed Lyme disease. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2008;47(7):910-4.
178. Wormser GP, Tang AT, Schimmoeller NR, Bittker S, Cooper D, Visintainer P, et al. *Utility of serodiagnostics designed for use in the United States for detection of Lyme borreliosis acquired in Europe and vice versa.* Medical microbiology and immunology. 2014;203(1):65-71.
179. Dubrey SW, Bhatia A, Woodham S, Rakowicz W. *Lyme disease in the United Kingdom.* Postgraduate Medical Journal. 2013;90(1059):33-42.
180. Leeflang MMG, Ang CW, Berkhout J, Bijlmer HA, Van Bortel W, Brandenburg AH, et al. *The diagnostic accuracy of serological tests for Lyme borreliosis in Europe: a systematic review and meta-analysis*. BMC infectious diseases. 2016;16(1):140.
181. Waddell LA, Greig J, Mascarenhas M, Harding S, Lindsay R, Ogden N. *The Accuracy of Diagnostic Tests for Lyme Disease in Humans, A Systematic Review and Meta-Analysis of North American Research.* PLOS ONE. 2016;11(12):e0168613.
182. Golightly MG. *Lyme borreliosis: laboratory considerations.* Seminars in neurology. 1997;17(1):11-7.
183. Rutkowski S, Busch DH, Huppertz HI. *Lymphocyte proliferation assay in response to Borrelia burgdorferi in patients with Lyme arthritis: analysis of lymphocyte subsets.* Rheumatology international. 1997;17(4):151-8.
184. Krause A, Brade V, Schoerner C, Solbach W, Kalden JR, Burmester GR. *T cell proliferation induced by Borrelia burgdorferi in patients with Lyme borreliosis. Autologous serum required for optimum stimulation.* Arthritis and rheumatism. 1991;34(4):393-402.
185. Zoschke DC, Skemp AA, Defosse DL. *Lymphoproliferative responses to Borrelia burgdorferi in Lyme disease. Annals of internal medicine.* 1991;114(4):285-9.
186. Valentine-Thon E, Ilsemann K, Sandkamp M. *A novel lymphocyte transformation test (LTT-MELISA) for Lyme borreliosis.* Diagn Microbiol Infect Dis. 2007;57(1):27-34.
187. Schempp C, Owsianowski M, Lange R, Gollnick H. *Comparison of Borrelia burgdorferi ultrasonicate and whole B, burgdorferi cells as a stimulus for T-cell proliferation and GM-CSF secretion in vitro.* Zentralblatt fur Bakteriologie : international journal of medical microbiology. 1993;279(3):417-25.
188. von Baehr V, Doebis C, Volk H-D, von Baehr R. *The Lymphocyte Transformation Test for Borrelia Detects Active Lyme Borreliosis and Verifies Effective Antibiotic Treatment.* The Open Neurology Journal. 2012;6:104-12.

189. van Gorkom T, Sankatsing SUC, Voet W, Ismail DM, Muilwijk RH, Salomons M, et al. *An Enzyme-Linked Immunosorbent Spot Assay Measuring Borrelia burgdorferi B31-Specific Interferon Gamma-Secreting T Cells Cannot Discriminate Active Lyme Neuroborreliosis from Past Lyme Borreliosis: a Prospective Study in the Netherlands.* J Clin Microbiol. 2018;56(4).
190. van de Schoor FR, Baarsma ME, Gauw SA, Joosten LAB, Kullberg BJ, van den Wijngaard CC, et al. *Validation of cellular tests for Lyme borreliosis (VICTORY) study*. BMC infectious diseases. 2019;19(1):732.
191. Klempner MS, Schmid CH, Hu L, Steere AC, Johnson G, McCloud B, et al. *Intralaboratory reliability of serologic and urine testing for Lyme disease.* Am J Med. 2001;110(3):217-9.
192. Magni R, Espina BH, Shah K, Lepene B, Mayuga C, Douglas TA, et al. *Application of Nanotrap technology for high sensitivity measurement of urinary outer surface protein A carboxyl-terminus domain in early stage Lyme borreliosis.* J Transl Med. 2015;13:346.
193. CDC. Lyme disease: Diagnostic testing. Available at: https://www.cdc.gov/lyme/resources/Halperin_2012_Chap4_JohnsonB.pdf [accessed 12/2018]
194. Schutzer SE, Angel TE, Liu T, Schepmoes AA, Clauss TR, Adkins JN, et al. *Distinct Cerebrospinal Fluid Proteomes Differentiate Post-Treatment Lyme Disease from Chronic Fatigue Syndrome.* PLOS One. 2011;6(2):e17287.
195. Liveris D, Schwartz I, Bittker S, Cooper D, Iyer R, Cox ME, et al. *Improving the Yield of Blood Cultures from Patients with Early Lyme Disease.* Journal of Clinical Microbiology. 2011;49(6):2166-8.
196. Soloski MJ, Crowder LA, Lahey LJ, Wagner CA, Robinson WH, Aucott JN. *Serum inflammatory mediators as markers of human Lyme disease activity.* PLOS One. 2014;9(4):e93243.
197. Aucott JN, Soloski MJ, Rebman AW, Crowder LA, Lahey LJ, Wagner CA, et al. *CCL19 as a Chemokine Risk Factor for Posttreatment Lyme Disease Syndrome: a Prospective Clinical Cohort Study*. Clinical and Vaccine Immunology. 2016;23(9):757.
198. Liveris D, Wang G, Girao G, Byrne DW, Nowakowski J, McKenna D, et al. *Quantitative detection of Borrelia burgdorferi in 2-millimeter skin samples of erythema migrans lesions: correlation of results with clinical and laboratory findings.* J Clin Microbiol. 2002;40(4):1249-53.
199. Mason LMK, Veerman CC, Geijtenbeek TBH, Hovius JWR. *Ménage à trois: Borrelia, dendritic cells, and tick saliva interactions.* Trends in Parasitology. 2014;30(2):95-103.
200. Johnson BJ, Pilgard MA, Russell TM. *Assessment of new culture method for detection of Borrelia species from serum of lyme disease patients*. J Clin Microbiol. 2014;52(3):721-4.
201. Kaforou M, Wright VJ, Oni T, French N, Anderson ST, Bangani N, et al. *Detection of Tuberculosis in HIV-Infected and -Uninfected African Adults Using Whole Blood RNA Expression Signatures: A Case-Control Study.* PLoS Med. 2013;10(10):e1001538.
202. Haas CT, Roe JK, Pollara G, Mehta M, Noursadeghi M. *Diagnostic 'omics' for active tuberculosis*. BMC Med. 2016;14:37.
203. Hytönen J, Kortela E, Waris M, Puustinen J, Salo J, Oksi J. *CXCL13 and neopterin concentrations in cerebrospinal fluid of patients with Lyme*

neuroborreliosis and other diseases that cause neuroinflammation. Journal of Neuroinflammation. 2014;11(1):103.
204. Hodzic E, Feng S, Holden K, Freet KJ, Barthold SW. *Persistence of Borrelia burgdorferi following antibiotic treatment in mice.* Antimicrob Agents Chemother. 2008;52(5):1728-36.
205. Jankova L, Chan C, Fung CLS, Song X, Kwun SY, Cowley MJ, et al. *Proteomic comparison of colorectal tumours and non-neoplastic mucosa from paired patient samples using iTRAQ mass spectrometry*. Molecular BioSystems. 2011;7(11):2997-3005.
206. Lebech AM, Hansen K, Brandrup F, Clemmensen O, Halkier-Sorensen L. *Diagnostic value of PCR for detection of Borrelia burgdorferi DNA in clinical specimens from patients with erythema migrans and Lyme neuroborreliosis*. Mol Diagn. 2000;5(2):139-50.
207. Anderson NL, Polanski M, Pieper R, Gatlin T, Tirumalai RS, Conrads TP, et al. *The Human Plasma Proteome: A Nonredundant List Developed by Combination of Four Separate Sources*. Molecular & Cellular Proteomics. 2004;3(4):311-26.
208. Lautner R, Zetterberg H, Blennow K. *Laboratory testing for lyme neuroborreliosis.* JAMA Neurol. 2015;72(1):125-6.
209. Molins CR, Ashton LV, Wormser GP, Hess AM, Delorey MJ, Mahapatra S, et al. *Development of a Metabolic Biosignature for Detection of Early Lyme Disease.* Clinical Infectious Diseases. 2015;60(12):1767-75.
210. Cheung CSF, Anderson KW, Benitez KYV, Soloski MJ, Aucott JN, Phinney KW, et al. *Quantification of Borrelia burgdorferi Membrane Proteins in Human Serum: A New Concept for Detection of Bacterial Infection.* Analytical Chemistry. 2015;87(22):11383-8.
211. Marques A, Schwartz I, Wormser GP, Wang Y, Hornung RL, Demirkale CY, et al. T*ranscriptome Assessment of Erythema Migrans Skin Lesions in Patients With Early Lyme Disease Reveals Predominant Interferon Signaling.* The Journal of infectious diseases. 2017;217(1):158-67.
212. Haas CT, Roe JK, Pollara G, Mehta M, Noursadeghi M. *Diagnostic 'omics' for active tuberculosis.* BMC medicine. 2016;14(1):37.
213. Montemurro F, Di Cosimo S, Arpino G. *Human epidermal growth factor receptor 2 (HER2)-positive and hormone receptor-positive breast cancer: new insights into molecular interactions and clinical implications.* Annals of Oncology. 2013.
214. Moniuszko A, Pancewicz S, Czupryna P, Grygorczuk S, Świerzbińska R, Kondrusik M, et al. *ssICAM-1, IL-21 and IL-23 in patients with tick borne encephalitis and neuroborreliosis.* Cytokine. 2012;60(2):468-72.
215. Molins CR, Ashton LV, Wormser GP, Hess AM, Delorey MJ, Mahapatra S, et al. *Development of a metabolic biosignature for detection of early Lyme disease. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America.* 2015;60(12):1767-75.
216. Pegalajar-Jurado A, Fitzgerald BL, Islam MN, Belisle JT, Wormser GP, Waller KS, et al. *Identification of Urine Metabolites as Biomarkers of Early Lyme Disease.* Scientific reports. 2018;8(1):12204.
217. Steere AC, Strle F, Wormser GP, Hu LT, Branda JA, Hovius JW, et al. *Lyme borreliosis.* Nature reviews Disease primers. 2016;2:16090.

218. Dinis-Oliveira RJ, Vieira DN, Magalhaes T. *Guidelines for Collection of Biological Samples for Clinical and Forensic Toxicological Analysis.* Forensic sciences research. 2016;1(1):42-51.
219. Hochmeister S, Engel O, Adzemovic MZ, Pekar T, Kendlbacher P, Zeitelhofer M, et al. *Lipocalin-2 as an Infection-Related Biomarker to Predict Clinical Outcome in Ischemic Stroke.* PLOS One. 2016;11(5):e0154797.
220. Kanda J, Mori K, Kawabata H, Kuwabara T, Mori KP, Imamaki H, et al. *An AKI biomarker lipocalin 2 in the blood derives from the kidney in renal injury but from neutrophils in normal and infected conditions.* Clinical and experimental nephrology. 2015;19(1):99-106.
221. Vergara A, Fernandez-Pittol MJ, Munoz-Mahamud E, Morata L, Bosch J, Vila J, et al. *Evaluation of Lipocalin-2 as a Biomarker of Periprosthetic Joint Infection.* The Journal of arthroplasty. 2019;34(1):123-5.
222. Flo TH, Smith KD, Sato S, Rodriguez DJ, Holmes MA, Strong RK, et al. *Lipocalin 2 mediates an innate immune response to bacterial infection by sequestrating iron.* Nature. 2004;432(7019):917-21.
223. Shao S, Cao T, Jin L, Li B, Fang H, Zhang J, et al. I*ncreased Lipocalin-2 Contributes to the Pathogenesis of Psoriasis by Modulating Neutrophil Chemotaxis and Cytokine Secretion.* The Journal of investigative dermatology. 2016;136(7):1418-28.
224. Srisawat N, Praditpornsilpa K, Patarakul K, Techapornrung M, Daraswang T, Sukmark T, et al. *Neutrophil Gelatinase Associated Lipocalin (NGAL) in Leptospirosis Acute Kidney Injury: A Multicenter Study in Thailand.* PLOS One. 2015;10(12):e0143367.
225. Petzke MM, Iyer R, Love AC, Spieler Z, Brooks A, Schwartz I. *Borrelia burgdorferi induces a type I interferon response during early stages of disseminated infection in mice.* BMC Microbiol. 2016;16:29-.
226. Hohl D, de Viragh PA, Amiguet-Barras F, Gibbs S, Backendorf C, Huber M. *The small proline-rich proteins constitute a multigene family of differentially regulated cornified cell envelope precursor proteins.* The Journal of investigative dermatology. 1995;104(6):902-9.
227. Hodge K, Have ST, Hutton L, Lamond AI. C*leaning up the masses: exclusion lists to reduce contamination with HPLC-MS/MS.* Journal of proteomics. 2013;88:92-103.
228. Cuypers B, Lecordier L, Meehan CJ, Van den Broeck F, Imamura H, Büscher P, et al. *Apolipoprotein L1 Variant Associated with Increased Susceptibility to Trypanosome Infection.* mBio. 2016;7(2):e02198-15.
229. Sun Y, Sheshadri N, Zong WX. S*ERPINB3 and B4: From biochemistry to biology.* Seminars in cell & developmental biology. 2017;62:170-7.
230. Geers C, Krüger D, Siffert W, Schmid A, Bruns W, Gro G. *Carbonic anhydrase in skeletal and cardiac muscle from rabbit and rat.* Biochem J. 1992;282 (Pt 1)(Pt 1):165-71.
231. Tan XT, Amran FB, Thayan R, Ahmad N, Jaafar R, Haron R, et al. *Potential serum biomarkers associated with mild and severe leptospirosis infection: A cohort study in the Malaysian population.* Electrophoresis. 2017;38(17):2141-9.
232. Tanaka T, Narazaki M, Kishimoto T. *IL-6 in inflammation, immunity, and disease.* Cold Spring Harb Perspect Biol. 2014;6(10):a016295-a.
233. Reis EAG, Hagan JE, Ribeiro GS, Teixeira-Carvalho A, Martins-Filho OA, Montgomery RR, et al. *Cytokine Response Signatures in Disease Progression and*

Development of Severe Clinical Outcomes for Leptospirosis. PLoS neglected tropical diseases. 2013;7(9):e2457.
234. Reis EA, Hagan JE, Ribeiro GS, Teixeira-Carvalho A, Martins-Filho OA, Montgomery RR, et al. *Cytokine response signatures in disease progression and development of severe clinical outcomes for leptospirosis.* PLoS neglected tropical diseases. 2013;7(9):e2457.
235. Moore A, Nelson C, Molins C, Mead P, Schriefer M. *Current Guidelines, Common Clinical Pitfalls, and Future Directions for Laboratory Diagnosis of Lyme Disease, United States.* Emerg Infect Dis. 2016;22(7).
236. Rebman AW, Crowder LA, Kirkpatrick A, Aucott JN. *Characteristics of seroconversion and implications for diagnosis of post-treatment Lyme disease syndrome: acute and convalescent serology among a prospective cohort of early Lyme disease patients*. Clinical rheumatology. 2015;34(3):585-9.
237. Klont F, Bras L, Wolters JC, Ongay S, Bischoff R, Halmos GB, et al. *Assessment of Sample Preparation Bias in Mass Spectrometry-Based Proteomics.* Analytical chemistry. 2018;90(8):5405-13.
238. Eaton SL, Roche SL, Llavero Hurtado M, Oldknow KJ, Farquharson C, Gillingwater TH, et al. *Total Protein Analysis as a Reliable Loading Control for Quantitative Fluorescent Western Blotting*. PLOS ONE. 2013;8(8):e72457.
239. Bass JJ, Wilkinson DJ, Rankin D, Phillips BE, Szewczyk NJ, Smith K, et al. *An overview of technical considerations for Western blotting applications to physiological research.* Scandinavian journal of medicine & science in sports. 2017;27(1):4-25.
240. Murphy RM, Watt KK, Cameron-Smith D, Gibbons CJ, Snow RJ. *Effects of creatine supplementation on housekeeping genes in human skeletal muscle using real-time RT-PCR*. Physiological genomics. 2003;12(2):163-74.
241. Castaño Z, Kypta R. *Housekeeping Proteins: Limitations as References During Neuronal Differentiation.* The Open Neuroscience Journal. 2008;2:36-40.
242. Dittmer A, Dittmer J. *Beta-actin is not a reliable loading control in Western blot analysis.* Electrophoresis. 2006;27(14):2844-5.
243. Fish-Low C-Y, Than LTL, Ling K-H, Lin Q, Sekawi Z. *Plasma proteome profiling reveals differentially expressed lipopolysaccharide-binding protein among leptospirosis patients*. Journal of Microbiology, Immunology and Infection. 2019.
244. Tape C, Tan R, Nesheim M, Kisilevsky R. *Direct evidence for circulating apoSAA as the precursor of tissue AA amyloid deposits*. Scandinavian journal of immunology. 1988;28(3):317-24.
245. Lybecker M, Henderson KC. *Borrelia burgdorferi Transcriptome Analysis by RNA-Sequencing.* Methods in molecular biology. 2018;1690:127-36.
246. Wu Q, Guan G, Liu Z, Li Y, Luo J, Yin H. *RNA-Seq-based analysis of changes in Borrelia burgdorferi gene expression linked to pathogenicity*. Parasit Vectors. 2015;8:155-.
247. Bouquet J, Soloski MJ, Swei A, Cheadle C, Federman S, Billaud J-N, et al. *Longitudinal Transcriptome Analysis Reveals a Sustained Differential Gene Expression Signature in Patients Treated for Acute Lyme Disease.* mBio. 2016;7(1).
248. Wang Z, Gerstein M, Snyder M. *RNA-Seq: a revolutionary tool for transcriptomics.* Nat Rev Genet. 2009;10(1):57-63.
249. Finotello F, Di Camillo B. *Measuring differential gene expression with RNA-seq: challenges and strategies for data analysis.* Briefings in Functional Genomics. 2014;14(2):130-42.

250. Cirillo E, Parnell LD, Evelo CT. *A Review of Pathway-Based Analysis Tools That Visualize Genetic Variants.* Front Genet. 2017;8:174-.
251. Frishberg A, Brodt A, Steuerman Y, Gat-Viks I. *ImmQuant: a user-friendly tool for inferring immune cell-type composition from gene-expression data.* Bioinformatics. 2016;32(24):3842-3.
252. Zaas AK, Burke T, Chen M, McClain M, Nicholson B, Veldman T, et al. *A host-based RT-PCR gene expression signature to identify acute respiratory viral infection.* Science translational medicine. 2013;5(203):203ra126.
253. Ravandi F, Walter RB, Freeman SD. *Evaluating measurable residual disease in acute myeloid leukemia.* Blood advances. 2018;2(11):1356-66.
254. Rauter C, Mueller M, Diterich I, Zeller S, Hassler D, Meergans T, et al. *Critical evaluation of urine-based PCR assay for diagnosis of Lyme borreliosis.* Clin Diagn Lab Immunol. 2005;12(8):910-7.
255. Brettschneider S, Bruckbauer H, Klugbauer N, Hofmann H. *Diagnostic value of PCR for detection of Borrelia burgdorferi in skin biopsy and urine samples from patients with skin borreliosis*. J Clin Microbiol. 1998;36(9):2658-65.
256. Carroll MW, Haldenby S, Rickett NY, Palyi B, Garcia-Dorival I, Liu X, et al. *Deep Sequencing of RNA from Blood and Oral Swab Samples Reveals the Presence of Nucleic Acid from a Number of Pathogens in Patients with Acute Ebola Virus Disease and Is Consistent with Bacterial Translocation across the Gut*. mSphere. 2017;2(4).
257. Shin H, Shannon CP, Fishbane N, Ruan J, Zhou M, Balshaw R, et al. *Variation in RNA-Seq transcriptome profiles of peripheral whole blood from healthy individuals with and without globin depletion.* PLOS One. 2014;9(3):e91041-e.
258. Bouquet J, Soloski MJ, Swei A, Cheadle C, Federman S, Billaud J-N, et al. *Longitudinal Transcriptome Analysis Reveals a Sustained Differential Gene Expression Signature in Patients Treated for Acute Lyme Disease.* mBio. 2016;7(1):e00100-16.
259. Frishberg A, Brodt A, Steuerman Y, Gat-Viks I. *ImmQuant: a user-friendly tool for inferring immune cell-type composition from gene-expression data.* Bioinformatics (Oxford, England). 2016;32(24):3842-3.
260. Shrestha N, Bahnan W, Wiley DJ, Barber G, Fields KA, Schesser K. *Eukaryotic initiation factor 2 (eIF2) signaling regulates proinflammatory cytokine expression and bacterial invasion.* The Journal of biological chemistry. 2012;287(34):28738-44.
261. Shrestha N, Bahnan W, Wiley DJ, Barber G, Fields KA, Schesser K. *Eukaryotic initiation factor 2 (eIF2) signaling regulates proinflammatory cytokine expression and bacterial invasion.* J Biol Chem. 2012;287(34):28738-44.
262. McCallum RE, Sword CP. *Mechanisms of pathogenesis in Listeria monocytogenes infection. VI. Oxidative phosphorylation in mouse liver mitochondria during experimental listeriosis.* Infect Immun. 1972;5(6):872-8.
263. Geyer S, Jacobs M, Hsu N-J. *Immunity Against Bacterial Infection of the Central Nervous System: An Astrocyte Perspective*. Fr. in Molecular Neuroscience. 2019;12:57.
264. Embers ME, Ramamoorthy R, Philipp MT. *Survival strategies of Borrelia burgdorferi, the etiologic agent of Lyme disease.* Microbes and infection. 2004;6(3):312-8.

265. Petnicki-Ocwieja T, Kern A. *Mechanisms of Borrelia burgdorferi internalization and intracellular innate immune signaling.* Fr. in cellular and infection microbiology. 2014;4:175-.
266. Lee SH, Dominguez R. *Regulation of actin cytoskeleton dynamics in cells.* Mol Cells. 2010;29(4):311-25.
267. Goldberg MB. *Actin-based motility of intracellular microbial pathogens*. Microbiol Mol Biol Rev. 2001;65(4):595-626.
268. Benedek O, Nagy G, Emody L. I*ntracellular signalling and cytoskeletal rearrangement involved in Yersinia pestis plasminogen activator (Pla) mediated HeLa cell invasion.* Micro. patho. 2004;37(1):47-54.
269. Troxell B, Xu H, Yang XF. *Borrelia burgdorferi, a pathogen that lacks iron, encodes manganese-dependent superoxide dismutase essential for resistance to streptonigrin*. J. Bio. Chem. 2012;287(23):19284-93.
270. Safire W. *How Not to Write.* 1st ed. New York Times Press. 1977:p82.

Appendix A: Ethical approval documents

NHS
Health Research Authority

East Midlands - Nottingham 2 Research Ethics Committee
The Old Chapel
Royal Standard Place
Nottingham
NG1 6FS

Please note: This is the favourable opinion of the REC only and does not allow you to start your study at NHS sites in England until you receive HRA Approval

26 July 2016

Mr Greg Joyner
PhD Student
PHE, University of Liverpool
IC2 Building
146 Brownlow Hill
Liverpool
L35RF

Dear Mr Joyner

Study title:	**Host-derived markers of Lyme disease and their diagnostic potential: A Discovery Study**
REC reference:	**16/EM/0330**
IRAS project ID:	**199388**

The Proportionate Review Sub-committee of the East Midlands - Nottingham 2 Research Ethics Committee reviewed the above application on 25 July 2016.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all studies that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact the REC Assistant Mrs Joanne O'Neil, NRESCommittee.EastMidlands-Nottingham2@nhs.net. Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the study.

Ethical opinion

On behalf of the Committee, the sub-committee gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation,

subject to the conditions specified below.

Conditions of the favourable opinion

The REC favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements. Each NHS organisation must confirm through the signing of agreements and/or other documents that it has given permission for the research to proceed (except where explicitly specified otherwise).

Guidance on applying for HRA Approval (England)/ NHS permission for research is available in the Integrated Research Application System, www.hra.nhs.uk or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of management permissions from host organisations.

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database. This should be before the first participant is recruited but no later than 6 weeks after recruitment of the first participant.

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to request a deferral for study registration within the required timeframe, they should contact hra.studyregistration@nhs.net. The expectation is that all clinical trials will be registered, however, in exceptional circumstances non registration may be permissible with prior agreement from the HRA. Guidance on where to register is provided on the HRA website.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion").

Summary of discussion at the meeting

The PR Sub-Committee confirmed the study raised no material ethical issues under the following headings: Social or scientific value; scientific design and conduct of the study, recruitment arrangements and access to health information, and fair participant selection, favourable risk benefit ratio; anticipated benefit/risks for research participants (present and future), care and protection of research participants; respect for potential and enrolled participants' welfare and dignity, informed consent process and the adequacy and completeness of participant information, suitability of the applicant and supporting staff, independent review, suitability of supporting information, other general comments and suitability of research summary.

Ethical issues raised, noted and resolved in discussion:

The PR Sub-Committee agreed that this was a well presented study with no material ethical issues.

The Sub-Committee commented that the application was not written in lay language and would have been difficult to understand if not for the lay summary which had to be asked for.

Approved documents

The documents reviewed and approved were:

Document	*Version*	*Date*
Covering letter on headed paper [Covering letter]		21 June 2016
Letter from sponsor [Declaration of PHE Sponsorship]		25 May 2016
Other [Julian Hiscox CV]	1	11 July 2016
Other [Nick Beeching CV]	1	11 July 2016
REC Application Form [REC_Form_15072016]		15 July 2016
Research protocol or project proposal [HOST-DERIVED MARKERS OF LYME DISEASE AND THEIR DIAGNOSTIC POTENTIAL: A DISCOVERY STUDY]	2	20 July 2016
Summary CV for Chief Investigator (CI) [GREG JOYNER CV]		26 May 2016
Summary CV for supervisor (student research) [Amanda Semper CV]		20 June 2016
Summary, synopsis or diagram (flowchart) of protocol in non technical language [LAY SUMMARY]	1	20 July 2016

Membership of the Proportionate Review Sub-Committee

The members of the Sub-Committee who took part in the review are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

Appendix A: Ethical Approval

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website:
http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at
http://www.hra.nhs.uk/hra-training/

With the Committee's best wishes for the success of this project.

16/EM/0330	**Please quote this number on all correspondence**

Yours sincerely

pp. S. O'Neil

Professor Frances Game
Chair

Email: NRESCommittee.EastMidlands-Nottingham2@nhs.net

Enclosures: *List of names and professions of members who took part in the review*

Copy to: *Dr Amanda Semper*
Mr Greg Joyner, PHE, University of Liverpool

Appendix A: Ethical Approval

NHS
Health Research Authority

NRES Committee London - Central
Skipton House
80 London Road
London
SE1 6LH

Telephone: 020 797 22560

10 April 2014

Dr Hilary Kirkbride
Consultant Epidemiologist
Public Health England
61 Colindale Avenue
London
NW9 5EQ

Dear Dr Kirkbride

Study title:	**Serum Archive for Emerging Zoonoses (SAfEZ): Establishment of a serum archive to assist in the risk assessment of potential new and emerging zoonoses.**
REC reference:	**14/LO/0308**
Protocol number:	**N/A**
IRAS project ID:	**137153**

Thank you for your letter of 17 March 2014, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information was considered by a sub-committee of the REC at a meeting held on 09 April 2014. A list of the sub-committee members is attached.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details, unless you expressly withhold permission to do so. Publication will be no earlier than three months from the date of this favourable opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to withhold permission to publish, please contact the REC Manager Mrs Audrey Adams, NRESCommittee.London-Central@nhs.net.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation

as revised, subject to the conditions specified below.

Ethical review of research sites

NHS sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Non-NHS sites

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database within 6 weeks of recruitment of the first participant (for medical device studies, within the timeline determined by the current registration and publication trees).

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non clinical trials this is not currently mandatory.

If a sponsor wishes to contest the need for registration they should contact Catherine Blewett (catherineblewett@nhs.net), the HRA does not, however, expect exceptions to be made. Guidance on where to register is provided within IRAS.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Advertisement	1.11	29 January 2014
Covering Letter		04 February 2014
Investigator CV	Hilary Kirkbride	01 August 2010
Letter from Sponsor		03 February 2014
Letter of invitation to participant	1.4	29 January 2014
Letter of invitation to participant	1.6	29 January 2014
Letter of invitation to participant	1.1	29 January 2014
Letter of invitation to participant	1.1	29 January 2014
Other: SAfEZ Acknowledgment Letter	1.4	23 October 2013
Other: SAfEZ Steering Group Terms of Reference	1.2	02 September 2013
Other: SAfEZ leaflet	1.12	12 March 2014
Participant Consent Form: SAfEZ	1.4	12 March 2014
Participant Information Sheet: AHVLA staff	1.7	10 March 2014
Participant Information Sheet: Bat workers	1.7	10 March 2014
Participant Information Sheet: Pig and Poultry fair 2014	1.2	10 March 2014
Participant Information Sheet: Sheep 2014 event	1.2	10 March 2014
Protocol	4.3	17 March 2014
Questionnaire: SAfEZ	1.5	12 March 2014
REC application	137153/5606 46/1/361	04 February 2014
Response to Request for Further Information	Letter from Dr Katherine Russell	17 March 2014

This Research Ethics Committee is an advisory committee to London Strategic Health Authority
The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document *"After ethical review – guidance for researchers"* gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

Feedback

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

Further information is available at National Research Ethics Service website > After Review

14/LO/0308	**Please quote this number on all correspondence**

We are pleased to welcome researchers and R & D staff at our NRES committee members' training days – see details at http://www.hra.nhs.uk/hra-training/

With the Committee's best wishes for the success of this project.

Yours sincerely

pp

Dr Andrew Hilson
Chair

This Research Ethics Committee is an advisory committee to London Strategic Health Authority
The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England

Appendix B: Presentation of Data

American Society of Microbiology Conference 2019

Congratulations! On behalf of the American Society of Microbiology and the ASM Microbe Sub-committee Reviewers, I am pleased to inform you that you have been selected to receive the **ASM Student and Postdoctoral Travel Award CPHM** to assist in covering your travel expenses to attend ASM Microbe 2018 in Atlanta, GA, on June 7-11, 2018.

Please confirm your acceptance no later than April 2, 2018. Please note that all recipients must be an ASM member to receive the award. To join ASM, please visit http://www.asmscience.org/content/membership/all. Additionally, please complete the Travel Award Form so we may process your award: https://asmeetings.formstack.com/forms/microbe2018_travel_award_form.

If you have not yet registered for ASM Microbe 2018 the discounted registration deadline is April 26, 2018. You can do register for the meeting here: https://www.asm.org/index.php/asm-microbe-2018-registration/2018-register.

Your check, in the amount of $750 will be available onsite, location to be determined. You will need to show photo ID in order to claim your check.

Please note: If you applied to more than one award, please note you were only accepted for this award.

If you applied to the following awards, you will receive a separate email with your acceptance or rejection notification.
1. Carlyn Halde Latin American Student Travel Award
2. ASM Undergraduate Research Capstone Program
3. ASM Undergraduate Research Fellowship
4. ASM New Tech Professional Development Award Program
5. NSF/ASM-LINK Undergraduate Faculty Research Initiative (UFRI) Fellowship

We look forward to hearing from you, and to your attendance at ASM Microbe 2018. Should you have any questions, please do not hesitate to contact ASM Staff by emailing asmmicrobemtgprogram@asmusa.org.

Sincerely,
ASM Meetings

American Society for Microbiology Conference, Atlanta 2018 – oral presentation abstract

Lyme disease (LD) is a multisystem infection caused by tick-borne spirochaetes of the Borrelia burgdorferi sensu lato group. UK laboratory diagnosis of Lyme disease involves the two-tier serological approach. The negative predictive value of the test has been challenged, particularly in early stage Lyme disease There is considerable interest, therefore, in the development of improved diagnostic tests. The main aim of the project is to identify new markers that could form the basis for improved tests. The project is part of collaboration between the University of Liverpool and Public Health England that aims to improve diagnostic testing for several infectious diseases.

A mass spectrometry biomarker discovery study was undertaken on Lyme disease positive and negative residual diagnostic samples from UK Lyme disease testing by Public Health England. A control group of healthy subjects serum samples (from NHS blood transfusion service) were also included. To ensure differences were specific to Lyme disease rather than genetic to infection, a "related-disease control group" including serum samples from syphilis, leptospirosis and chronic fatigue syndrome were included. A total of 50 human samples were compared by label-free quantitative mass spectrometry.

Surprisingly, Lyme disease seropositive and seronegative groups were found to have very few proteins that were significantly different when directly compared. One protein, Lipocalin-2 was found at a significantly higher abundance in the LD-positive patients compared with those that were LD-negative. This is of interest due to involvement in innate immunity. Lipocalin-2 has been found in mice exposed to B. burgdorferi. Leptospirosis samples showed increased levels of several proteins involved in host-immune response including neutrophil defensin-1. Several key differences were also found in CFS and syphilis patients.

The results of the mass spectrometry run have generated several proteins of interest that will be further investigated by Western blot analyses on larger sample groups to further investigate their diagnostic potential.

Oral and Poster presentations

Microbiology Society Conference 2016, Liverpool	21-24 March 2016	Poster presentation: "Lyme disease: Proteomics and discovery of potential biomarkers"
Liverpool Brain Infections Group Retreat 2016	14-16 December 2017	Oral presentation: "Lyme neuroborreliosis, background and diagnostics"
HPRU annual scientific conference, Manchester	06-07 March 2017	Oral presentation "Lyme disease: Proteomics-based discovery of potential biomarkers:
American Society for Microbiology Conference, Atlanta 2018	08-10 June 2018	Oral presentation. Awarded travel grant to attend. "Host-derived markers of Lyme disease and how to find them"
Department of Infection Biology fortnightly seminar	09 October 2018	Oral presentation: "Host-derived markers of Lyme disease and other spirochaetal infections: Discovery and diagnostic potential"
HPRU annual scientific conference	12-13 December 2019	Oral presentation "Lyme disease: Issues with diagnostics and how can could be improved"
Microbiology Society Conference 2019, Belfast	08-11 April 2019	Poster Presentation "Host Biomarkers of Lyme Disease"

Example poster

NIHR Health Protection Research Unit in Emerging and Zoonotic Infections

Host-derived biomarkers: Their discovery and diagnostic potential in Lyme disease and other spirochaetel infections

Greg P Joyner (PhD Student) Institute of Infection and Global Health, University of Liverpool / Public Health England Porton
Supervisors: **Dr. Nick Beeching**, LSTM; **Prof. Julian Hiscox**, University of Liverpool; **Dr. Amanda Semper** and **Dr. Tim Brooks**, PHE Porton

1 Introduction

Background: Lyme disease (LD) is a multisystem infection caused by tick-borne spirochaetes of the *Borrelia burgdorferi* sensu lato group. **Human cases have increased steadily** in the UK in recent years with over 1100 confirmed cases in 2013. In the same year, over 27,000 cases were reported in the US. **UK laboratory diagnosis of LD involves the internationally accepted two-tier serological approach**: a C6 ELISA followed by confirmatory IgG and IgM line blots. **The negative predictive value of the test has been challenged**, particularly in early stage LD and in some chronic infections. **There is considerable interest, therefore, in the development of improved diagnostic tests**.

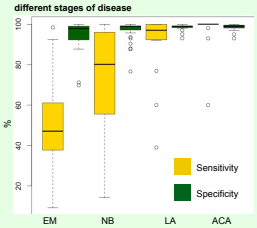
Aims:
- Review of current UK testing protocols
- Discovery of host-derived markers of early LD and potential risk markers for development of long term symptoms or failed response to treatment.

Hypotheses:
- That novel host biomarkers are characteristic of LD and have diagnostic potential
- That host biomarkers present early the course of infection are of prognostic value in determining the clinical manifestation or outcome of LD

2 Diagnostic test review

Analysis of current literature on sensitivity and specificity of serological tests for LD showed lower sensitivity/specificity at early stages of disease (EM). The development of antibodies in response to infection, particularly of IgG, may take over 4-6 weeks to reach detectable levels. Results from studies investigating the efficacy of Lyme serology were highly variable giving broad confidence intervals. The data also suggested that two-tiered tests did not outperform single tests at all stages of disease. However, results were highly variable with many studies showing high risk of bias

Sensitivity and specificity of LD serology (TTTA) at different stages of disease



Right: Sensitivity and specificity of LD serology at different stages of disease
Total patients: Acute stage (Erythema migrans rash) n=3758, neuroborreliosis n=1120, Lyme arthritis n=458, Acrodermatitis chronica atrophicans n=462
EM = Erythema migrans (acute stage), NB = neuroborreliosis, LA = Lyme arthritis, ACA = Acrodermatitis chronica atrophicans

3 Biomarker discovery protocol

Residual serum samples from UK LD diagnostic testing at RIPL, PHE Porton
Positive n=15
Negative n=15

Control patient samples

Control samples from patients with related-diseases (similar morphology of causative agent or overlapping clinical symptoms)

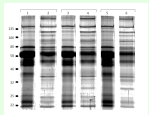
Healthy n=5
Syphilis n=5
Leptospirosis n=5
CFS n=5

(i) Serum depletion: Top-12 spin columns were used to remove several proteins present in blood at high concentrations including serum albumin and IgG

(ii) Quantitative mass spectrometry: Discovery of proteins which show significant changes in abundance consistently between the two groups

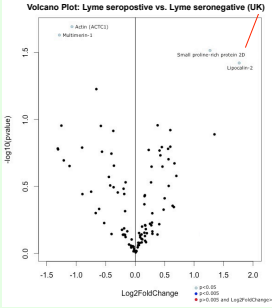
(iii) Biomarker validation
- Can protein biomarkers differentiate between patients with (seropositive) and without (seronegative, healthy) LD
- Western blot analysis – using antibodies for proteins discovered in discovery mass spectrometry run.

The SYPRO ruby stained electrophoresis gel (right) shows protein bands before and after depletion with equal quantities of protein loaded (2.5µg). Lanes 1,3,5 show serum before depletion with 2,4,6 showing samples after depletion.



4 Mass spectrometry results

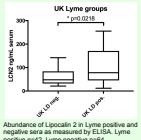
Lyme disease



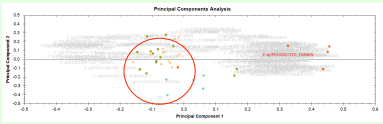
Volcano plot showing proteins significantly increased (right) or decreased (left) in Lyme disease positive sera when compared to Lyme negative sera

Lipocalin-2 (LCN2): Protein involved in innate immunity expressed in neutrophils. Despite presence in many infections, presence of lipocalin-2 was shown to to be significantly increased in abundance in LD positive samples.

Fewer differences between Lyme seropositive and seronegative were found than expected, with both groups closely grouping together in PCA.



Abundance of Lipocalin 2 in Lyme positive and negative sera as measured by ELISA. Lyme positive n=42. Lyme negative n=64

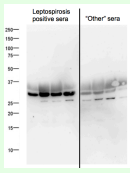


PCA (principal component analysis) of mass spectrometry data. Red circle shows grouping of Lyme positive and Lyme negative sera

There are several possible reasons for the lack of
- Sample quality: Samples used in initial MS run were residual diagnostic samples with unknown storage history. Repeated freeze-thaw and poor storage may have led to degradation of samples.
- True result: *Borrelia* is only transiently present in the blood, the differences in patients with LD may be too subtle to detect using current proteomic techniques

These results will be further investigated using further MS and transcriptomic analysis

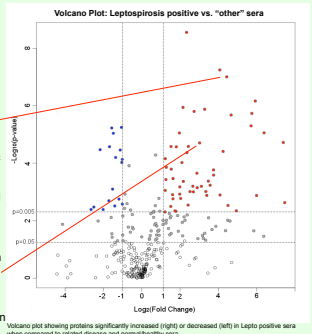
Leptospirosis: In contrast to the Lyme disease sera, leptospirosis positive sera was found to be highly distinct to other sera and grouped closely based on the proteome of individual samples. Proteins known to be elevated in leptospirosis infection, including the current biomarker Creatine Kinase M, were identified as significantly increased in abundance in the sample set. Further proteins of interest were identified for follow-up analysis by WB and ELISA.



WB for CA3: Left: sera from Lepto positive patients (not included in MS). Right: other sera from related diseases and normal/ healthy patients

Carbonic anhydrase III (CA3): catalyzes the reversible hydration of carbon dioxide. CA3 is highly tissue specific and present in high levels in skeletal muscles and at lower levels in cardiac and smooth muscle. It's inreased presence in leptospirosis patient sera may be due to rhabdomyolysis (muscle breakdown during infection)

Myosin (heavy chain 7): Actin motor protein with role in muscle contraction. Potential marker of myocardial cell damage but is also found in skeletal muscles and high serum levels may indicate muscle trauma due to infection



Volcano plot showing proteins significantly increased (right) or decreased (left) in Lepto positive sera when compared to related disease and normal/healthy sera

5 Future work

- **Further quantification** of potential biomarkers of Lyme and leptospirosis using WB and ELISA using serum bank
- **Mass Spec II:** In order to investigate the lack of differences between LD seropositive and seronegative, a second mass spectrometry run will use well-characterised, freshly collected sera
- **Transcriptomic analysis:** Whole-blood samples are currently being collected for use in a transcriptomic-based study of the host-response to early Lyme disease. Samples are being collected on presentation, after antibiotic treatment and then at 3 monthly intervals up to 1 year. Data will be analysed on MinION and Illumina RNAseq platforms and compared to normal/ healthy volunteer blood. Differential gene analysis will be used with the aim of identifiying a minimal marker set that is able to distinguish early Lyme samples

LSTM LIVERPOOL SCHOOL OF TROPICAL MEDICINE
UNIVERSITY OF LIVERPOOL | INSTITUTE OF INFECTION AND GLOBAL HEALTH
Public Health England
Funded by NHS National Institute for Health Research

The research was funded by the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Emerging and Zoonotic Infections at University of Liverpool in partnership with Public Health England (PHE), in collaboration with Liverpool School of Tropical Medicine. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health or Public Health England.